European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Health insurance, portfolio choice, and retirement incentives

Emilio Barucci[a], Enrico Biffis[b], Daniele Marazzina[a,*]

[a] Department of Mathematics, Politecnico di Milano, I-20133 Milano, Italy
[b] Imperial College Business School, Imperial College London, South Kensington Campus, SW7 2AZ United Kingdom

ARTICLE INFO

Article history:
Received 15 April 2022
Accepted 19 September 2022
Available online xxx

Keywords:
Finance
Health insurance
Health shock
Retirement
Optimal asset allocation

ABSTRACT

We study optimal portfolio choice and labor market participation in a continuous time setting in which agents face health shocks, medical expenses, and random lifetimes. We explore the implications of different forms of health coverage and study their impact on dynamic portfolios and labor supply decisions. We characterize these effects in semi-closed form, providing tools to measure retirement incentives as a function of relevant state variables and health cover arrangements. A calibration of the model matches empirically observed labor market participation patterns and portfolio decisions of US workers during the last phase of their working lives, while offering insights into the interlinkage between labor market participation, health insurance provision and portfolio choice.

1. Introduction

For many individuals the aspiration to retire is an important driver of wealth planning and life style choice. Improvements in life expectancy and rising health care costs, however, make the prospect of outliving one's resources increasingly material. These considerations affect not only investment decisions during the accumulation and decumulation phases of the life cycle, but also the choice of when to retire. Governments have long recognized that the secular increase in life expectancy and fall in interest rates pose severe challenges for the sustainability of social security and pension provision. Although the switch from defined benefit to defined contribution retirement plans has partly mitigated the burden of pension liabilities,[1] a number of policy questions remain open, such as whether to increase the minimum retirement age or how to design eligibility criteria for medical coverage. These questions can be properly addressed only by taking into account the fiscal effects of policy intervention, which in turn cannot abstract away from labor market participation incentives.

In this paper we provide a parsimonious continuous time model capturing some of the main trade offs driving investment and labor market participation decisions during the latter part of an agent's working life, as well as portfolio choice after retirement. In particular, we focus on the retirement (dis)incentives of medical expenses and various insurance mechanisms ranging from private insurance to medical coverage offered by employers and governments (e.g., Medicare in the US). We consider individuals facing a random life time whose distribution is shaped by health shocks leading in turn to medical expenses and potentially lower earnings. In the baseline model, agents can self-insure by investing in a riskless and a risky asset, or can insure themselves via life and health covers available in the private insurance market. They can also adjust continuously their labor supply as well as make an irreversible retirement decision associated with a jump in leisure. We then add employer and government provided health insurance: we consider common configurations offering health coverage while working and/or after retirement.

Throughout the first part of our analysis, we consider a complete market in which all sources of risk are insurable. In an extension of the model, we consider an unspanned source of risk, which can be thought of as an uninsurable health shock leading to a reduction in life expectancy and higher future medical expenses. The results are shown to be robust to such an extension, which requires additional technical work. The baseline model is therefore simple enough to be solved in semi-closed form, yet rich enough to deliver empirical predictions in line with recent evidence on medical expenses and labor market participation (e.g., French, 2005; French & Jones, 2011), as well as portfolio choice among the elderly (e.g., De Nardi, French, & Jones, 2009; 2010). The tractability of the model makes it suitable for comparative

* Corresponding author.
E-mail addresses: emilio.barucci@polimi.it (E. Barucci), e.biffis@imperial.ac.uk (E. Biffis), daniele.marazzina@polimi.it (D. Marazzina).

[1] See Maurer, Mitchell, Warshawsky, & Warshawsky (2012) for an overview of some of these issues.

https://doi.org/10.1016/j.ejor.2022.09.016
0377-2217/

statics analysis in the context of health care and social security reforms.

The findings of the paper address two main research questions: i) gauging the (dis)incentives for labor market participation provided by different forms of health insurance schemes, and ii) understanding the optimal portfolio choice of an agent that can modulate investment, insurance decisions, and labor market participation in response to insurable and uninsurable health shocks.

To answer the first question, we follow French & Jones (2011) and consider the following insurance frameworks: private insurance only ("private insurance" case); employer-provided insurance offering coverage only while working ("tied insurance" case); employer-provided insurance offering coverage while working as well as during retirement ("retiree insurance" case); coverage provided only during retirement, which we regard as a proxy for Medicare in the US and similar arrangements elsewhere ("Medicare" case). We find that the private and tied insurance cases provide the greatest incentives to delay retirement, whereas retiree insurance and Medicare incentivize earlier retirement to gain access to health coverage. The optimal portfolio strategies in the Medicare and private insurance case are rather similar, but clearly agents retire much earlier on average with Medicare, as health insurance does not need to be funded during retirement and hence wealth decumulation can be slower. Irrespective of the type of cover considered, we observe decreasing patterns in risky asset demand ahead of retirement. The introduction of an unspanned source of health risk amplifies these effects and makes clearer the divergent role of life expectancy and medical expenses in shaping labor market participation. On the one hand, a reduction in life expectancy allows agents to revise downwards their "distance to retirement" thus disincentivizing labor market participation. On the other hand, health shocks leads to immediate medical expenses or an increase in future medical expenses, thus providing an incentive to generate wages to support immediate and future health costs. We find that the second effect dominates in the private and tied insurance cases, making the distance to retirement increase, whereas the first effect dominates in the Medicare and retiree insurance cases, for which retirement is instead accelerated. The option-like element introduced by irreversible retirement (see discussion below) leads to interesting nonlinear effects for a number of model parameters. The most interesting insights we obtain are probably associated with the likelihood of health shocks, which can have opposite effects on the incentives to retire, depending on the particular health insurance framework considered. For example, an increase in the likelihood of health deterioration has opposite implications for Medicare and tied insurance, making the distance to retirement reduce in the former case and increase in the latter.

As far as the second question is concerned, we show that during their working life agents adjust continuously labor supply in response to wage levels and their preferences for leisure and consumption, while keeping track of a wealth dependent threshold triggering irreversible retirement once reached. This is in line with Farhi & Panageas (2007), but in our setup the wealth threshold, and hence the distance to retirement, is health dependent: the better the health state of the agent, the greater the distance to retirement, as more sizeable resources need to be accumulated to support consumption and rising medical expenses over a longer lifespan. Our model is flexible enough to deliver a rich set of empirical predictions consistent with more or less aggressive investment strategies ahead of retirement. In particular, in our estimation of the model based on data from the Health and Retirement Study (HRS), we find that the risky asset allocation decreases relative to total wealth as the individual approaches retirement. Importantly, portfolio choice depends not only on the current health state, but also on the possibility of future health transitions and associated medical expenses, which leads to less aggressive risky asset allocation as the retirement goal approaches. The same pattern applies to private health insurance demand, which is initially large but then declines on average as the retirement goal becomes closer and excessive insurance purchases would prevent wealth from accumulating fast enough.

The paper is organized as follows. The next section discusses the existing literature. Section 2 introduces the setup and presents the baseline model. In Section 3, we give an idea of how the model is solved and discuss the solutions for optimal labor supply, consumption, and investment/insurance decisions. Section 4 presents results based on estimation of the model based on HRS data. We consider first the case in which only private health insurance is available and then introduce the availability of employer-provided insurance. We then consider different health insurance configurations and carry out sensitivity analyses of the optimal retirement threshold relative to key parameters of interest. Section 5 introduces an unpsanned source of risk. Finally, Section 6 offers some concluding remarks. We relegate to the online supplementary material the proofs of the main results of the paper, as well as a number of closed and semi-closed form expressions for key quantities of interest.

1.1. Literature review

The paper is related to at least two strands of literature. First, the paper contributes to the vast literature on lifecycle portfolio choice originating from the seminal contributions of Samuelson (1969) and Merton (1971). The paper speaks in particular to the line research focusing on endogenous labor market participation. An important reference in this area is the work of Bodie, Merton, & Samuelson (1992), who allow agents to adjust labor supply continuously. The empirical evidence, however, suggests that the latter is to a large extent indivisible, as many workers who retire do not return to work at a later date and, if they do, they work only part time or for lower wages (e.g., Hausman & Wise, 2008). Farhi & Panageas (2007) and Dybvig & Liu (2010) allow for irreversibility of the retirement decision, demonstrating how this introduces nonlinear option-like effects in the agents' optimal strategies. In particular, Farhi & Panageas (2007) show how an agent's wealth plays a dual role, as it determines not only the resources available for future consumption, but also the distance to retirement, as the agent retires only when reaching a high enough wealth threshold. This aspect is not material when labor supply can only be adjusted continuously. In line with Choi, Shim, & Shin (2008), we consider both continuous and irreversible labor supply adjustments. In addition to the extant literature, we explicitly allow for health risks to shape the agent's life expectancy (see Hugonnier, Pelgrin, & St-Amour, 2013) as well as trigger medical expenses and lower wages (productivity losses). On the methodological side, we solve the continuous time optimal portfolio problem by using a duality approach, in the spirit of He & Pages (1993) and Karatzas & Wang (2000). Papers that are closely related to ours are Farhi & Panageas (2007), Choi et al. (2008), Dybvig & Liu (2010), Barucci & Marazzina (2011), Bensoussan, Jang, & Park (2016), Chai, Horneff, Maurer, & Mitchell (2011). Differently from our contribution, they abstract away from insurance and risky wages (Choi et al., 2008; Farhi & Panageas, 2007), labor flexibility before the irreversible retirement decision (Barucci & Marazzina, 2011; Dybvig & Liu, 2010), and from health shocks and insurance (Bensoussan et al., 2016; Chai et al., 2011). We also refer to Koo, Pantelous, & Wang (2022) for a related discrete time life cycle model, and to Konicz & Mulvey (2015) and Owadally, Jang, & Clare (2021) for a multi-stage stochastic programming approach.

The second strand of literature our contribution speaks to focuses on the impact of medical expenses on labor market participation and portfolio choice. The importance of social security and health insurance provisions in this context has been studied by a number of authors, including Rust & Phelan (1997), Blau & Gilleskie (2008), Hubbard, Skinner, & Zeldes (1995), French (2005), Hausman & Wise (2008), Blundell, French, & Tetlow (2016), French & Jones (2011). Earlier contributions such as Lumsdaine, Stock, & Wise (1996) and Gustman & Steinmeier (1994) consider medical expenses and employer-provided health insurance finding modest impact on labor market participation, but this is largely due to assuming that health insurance has no role other than reducing average medical expenses. The dynamic programming models estimated by Rust & Phelan (1997) and Blau & Gilleskie (2008) show instead that when agents are risk averse and health insurance helps mitigate the volatility of out-of-pocket medical expenses, then labor supply responses can be much larger. French (2005) and French & Jones (2011) take this point further by allowing individuals to save to smooth consumption and self-insure against volatile medical expenses. French & Jones (2011) obtain a good match of the empirical evidence on 60-year-old US males covered by HRS and consider different health insurance schemes to estimate their impact on labor supply. Their model is in discrete time and the dynamic programming problem is solved numerically. We consider a similar model, but in continuous time, and use a duality approach to reduce its solution to solving a system of free boundary problems in a regime switching framework, thus delivering semi-analytical expressions for a number of quantities of interest. We obtain results consistent with those documented in French & Jones (2011), but are also able to explain the mechanics of labor market participation decisions induced by the option-like nature of irreversible retirement, in the spirit of Farhi & Panageas (2007) and Chai et al. (2011). Another important question is how individuals save during retirement. The empirical evidence suggests that the elderly consume more frugally than standard lifecycle models would predict, a possible reason being rising life expectancy and medical expenses (e.g., De Nardi et al., 2009; 2010; Dynan, Skinner, & Zeldes, 2004). Our model and its empirical calibration provide support for these insights and show how medical expenses associated with health deterioration have a larger impact on retirement decisions than life expectancy considerations. An interesting question, which is not addressed in this paper, is how health investment would impact labor supply decisions. Hugonnier et al. (2013) propose a continuous time portfolio choice model with endogenous health risk which can be related to ours in the way health dynamics shape life expectancy. Although they do not consider flexibility in labor supply and irreversible retirement, they obtain important insights into the wealth dependence of health expenditures. The exploration of this angle in the context of labor supply responses is left for future research.

2. The setup

We consider an agent with initial wealth $W(0) > 0$ and an endowment of leisure normalized to one unit. At each time $t \geq 0$ the agent chooses consumption $c(t)$, leisure $l(t)$ (equivalently, labor supply $1 - l(t)$), as well as how to allocate her wealth $W(t)$ to a riskless and a risky asset. Labor market participation generates income $Y(t)$. The agent can decide to exit the labor market once and for all at the endogenous retirement date τ_r. Life expectancy evolves over time in response to health shocks, which trigger random medical expenses. Life and health insurance are available throughout the agent's lifetime. We therefore work in partial equilibrium, in the spirit of Merton (1971), as the agent takes wages, health dynamics and asset prices as given.

In the following, we outline a simple model allowing us to determine semi-explicitly the agent's optimal portfolio choice and labor market participation decisions. An extension of the baseline model to a richer setting including longevity or pandemic risk and multiple health states is discussed in Section 5. As notation can be rather daunting at times, we facilitate navigation of the setup by using subscripts "d", "h" and "S" for quantities related to **d**eath, **h**ealth, and **s**tocks (risky assets), respectively.

2.1. Health and mortality shocks

The agent's planning horizon is bounded by the random death time τ_d, which coincides with the first jump of a conditionally Poisson process[2] with intensity $\lambda_d(H(t)) > 0$. The latter represents the conditional instantaneous death probability, given the agent's health state, $H(t)$, prevailing at time t. In the simplest model specification, we consider the two states "best" and "poor", so that we have $H(t) \in \{\mathrm{b}, \mathrm{p}\}$. An agent in the best health state has intensity of mortality $\lambda > 0$ (i.e., $\lambda_d(\mathrm{b}) = \lambda$). When a health shock occurs at an independent Poisson time τ_h with parameter $\lambda_h > 0$, the agent transitions to the poor health state and the intensity of mortality jumps to level $\lambda_d(\mathrm{p}) = \lambda + \Delta_h$, $\Delta_h \geq 0$. For simplicity, we assume health state transitions to be irreversible. We therefore have that the intensity of mortality satisfies

$$\lambda_d(t) := \lambda_d(H(t)) = \lambda + \Delta_h 1_{\tau_h \leq t}. \tag{2.1}$$

In line with Markov chain models of health dynamics, one may regard death occurrence as a transition to a third and absorbing state (e.g., Asmussen & Steffensen, 2020; Hoem, 1969). We note that the model can be extended to any finite number of health states; see Section 5 for an explicit example and Chen, Chang, Sun, & Yu (2022) for a formulation allowing the dynamics of the health status to be driven by a Brownian motion. The use of two health states clearly allows us to obtain neater solutions, while capturing the most salient features of the data.

In addition to reducing life expectancy, the health shock results in medical expenses amounting to the random quantity $M \geq 0$ a.s.. We assume M to be an independent, square integrable random variable capturing treatment costs. The health shock also induces productivity losses captured by lower hourly wages and reduced leisure endowment as discussed in the next section.

Extension of the model to medical expenses dependent on the health state is straightforward and is illustrated in Section 5.

2.2. Labor income and leisure

The agent is endowed with one unit of leisure, a portion $(1 - l(t))$ of which can be allocated to work, which delivers in turn a flow of income

$$Y(l(t), w(H(t))) = (1 - l(t)^p) w(H(t)), \tag{2.2}$$

with $p \geq 1$ and $0 \leq l(t) \leq \bar{l}(H(t)) < 1$. In the above, $w(H(t))$ denotes the wage rate, which is allowed to depend on the health status, as in French (2005) and French & Jones (2011). The idea here is for our stylized model to capture in reduced form the fact that the wage rate experienced by healthy and unhealthy individuals might differ. A health shock is likely to reduce worker's productivity which leads to a reduction of leisure endowment and, in some cases, of the wage rate, the phenomenon being particularly relevant in case of a self-employed worker (see Grossman, 1972; Lee, 1982, for a foundation of the relationship between wage and health status handling them as endogenous). The model specification also allows us to match the empirical findings of French (2005) for wages of individuals aged between 50 and 60, which are clearly relevant for our analysis. In line with Aaronson & French (2004), the parameter p allows us to capture the empirical regularity that, all else equal, part-time workers earn relatively higher

[2] See, for example, Biffis, Denuit, & Devolder (2010).

wages than full time workers. Aaronson & French (2004) find that a 50% drop in work hours leads to a 25% drop in the offered hourly wage. The quantity $\bar{l}(H(t))$ provides a cap on leisure that can be enjoyed before retirement, and captures the fact that the agent may have to spend time training or looking for a job when unemployed; we may interpret the quantity $(1-\bar{l}(H(t))^p)w(H(t))$ as unemployment benefits. In line with Choi et al. (2008) and Barucci & Marazzina (2011), the cap also ensures that the retirement time is finite, for otherwise the agent would be able to enjoy full leisure and never retire. We assume $\bar{l}(j)=\bar{l}-(\bar{l}-\underline{l})1_{j=\mathrm{p}}$, with $j\in\{\mathrm{b},\mathrm{p}\}$ and $\bar{l}-\underline{l}>0$, to account for the fact that the agent might experience a further reduction in leisure after a health shock occurs. Although wages could in principle vary over the life cycle, in the baseline model we allow them to change in response to the health shock only, and simply write $w(j)=\overline{w}-(\overline{w}-\underline{w})1_{j=\mathrm{p}}$, with $j\in\{\mathrm{b},\mathrm{p}\}$ and $\overline{w}-\underline{w}\geq 0$.

Retirement corresponds to full leisure and zero labor income, i.e., $l(t)=1\ \forall t>\tau_r$, where τ_r is the retirement time. The retirement decision is therefore irreversible.

2.3. Investment opportunity set

The agent can invest in a money market account, which pays continuously the riskless rate $r>0$, and in a risky asset with gain process S evolving according to

$$\mathrm{d}S(t)=S(t)(b\mathrm{d}t+\sigma\mathrm{d}Z(t)),\quad S(0)=S_0, \tag{2.3}$$

where $b,\sigma>0$ are given constants, and Z is a standard Brownian motion. The gain process represents the value of a portfolio that continually reinvests any dividends paid out by the risky asset.

There is also a private market to insure against medical expenses and mortality risk at actuarially fair prices:

- Life insurance and annuities. By paying a premium $\lambda_d(t)(\theta_d(t)-W(t))$ at time t, the agent ensures that her beneficiaries receive a death benefit equal to $\theta_d(t)-W(t)$ should death occur over the next small time interval. Here, $\theta_d(t)-W(t)$ is the face value of the life insurance contract and is chosen by the agent via the bequest target $\theta_d(t)$. As in Dybvig & Liu (2010), we interpret the contract as an annuity whenever $\theta_d(t)<W(t)$.
- Health insurance. In line with Koijen, Van Nieuwerburgh, & Yogo (2016), the agent may have access to employer-provided coverage while working and Medicare in retirement. We denote by $\eta(t)$ the fraction of health expenses covered by such arrangements.[3] The agent has also access to supplemental, private health insurance. By paying a premium $\theta_h(t)\lambda_h E[M]$ at time t, the agent has the right to receive the amount $\theta_h(t)M$ in case the health shock occurs over the next small time interval. Full insurance is delivered by a choice of $\theta_h(t)$ equal to $1-\eta(t)$.

2.4. The agent's optimization problem

At each time t before death, the agent chooses the wealth amount allocated to the risky stock, $\theta_S(t)$, the bequest target delivered by life insurance, $\theta_d(t)$, and the face value of health insurance, $\theta_h(t)E[M]$ before the health shock. The agent also decides consumption, $c(t)\geq 0$, leisure allocation, $l(t)\geq 0$, as well as the retirement date τ_r. To make precise the optimization problem and admissible strategies, we work on the filtered probability space $(\Omega,\mathcal{G},\mathbb{G},\mathbb{P})$, where the filtration $\mathbb{G}:=(\mathcal{G}_t)_{t\geq 0}$ is such that each sigma-field $\mathcal{G}_t$ is defined as $\mathcal{G}_t:=\cap_{u>t}\mathcal{F}_t\vee\sigma_g(\tau_d\wedge u)$, where $\sigma_g(U)$ denotes the sigma-field generated by the random variable U. Each sigma-field $\mathcal{F}_t$ is in turn defined as $\mathcal{F}_t:=\cap_{u>t}\mathcal{F}^Z_t\vee\sigma_g(\tau_h\wedge u)$, where the filtration $\mathbb{F}^Z:=(\mathcal{F}^Z_t)_{t\geq 0}$ is the one generated by the Brownian motion Z and augmented with the $\mathbb{P}$-null sets. This construction makes $\mathbb{G}$ the smallest enlargement of the Brownian filtration ensuring that the random times τ_d and τ_h are stopping times (e.g., Protter, 2005, Section VI.3, page 370) and will be shown to be most useful when simplifying the optimization problem (2.7) further below.

A portfolio strategy $(c,l,\boldsymbol{\theta},\tau_r)$, with $\boldsymbol{\theta}:=(\theta_S,\theta_h,\theta_d)$, is admissible if it is $\mathbb{G}$-predictable and is such that c,l,θ_h,θ_d are integrable and θ_S is square-integrable. In this framework, the agent's wealth process satisfies the dynamic budget constraint

$$\begin{aligned}\mathrm{d}W(t)=&(1-N_d(t))\Big\{[Y(l(t),w(H(t)))-c(t)]\mathrm{d}t\\&+(W(t)-\theta_S(t))r\mathrm{d}t\\&+\theta_S(t)(b\mathrm{d}t+\sigma\mathrm{d}Z(t))-\lambda_d(t)(\theta_d(t)-W(t))\mathrm{d}t\\&-(1-N_h(t))\lambda_h\theta_h(t)E[M]\mathrm{d}t\\&-M(1-\eta(t)-\theta_h(t))\mathrm{d}N_h(t)\Big\}\\&+(\theta_d(t)-W(t-))\mathrm{d}N_d(t),\end{aligned} \tag{2.4}$$

where we denote by $N_i(t):=1_{\tau_i\leq t}$ the indicator of the intensity jump at time τ_i, with $i\in\{h,d\}$. The agent maximizes her lifetime expected utility from leisure and consumption. We assume time separable preferences with subjective discount rate $\delta>0$, and define the agent's utility flow as

$$u_c(c(t),l(t))=\frac{\left(l(t)^{1-\alpha}c(t)^{\alpha}\right)^{1-\gamma}}{1-\gamma},\quad \gamma>0,\ 0<\alpha<1. \tag{2.5}$$

A value $\gamma<1$ or $\gamma>1$ means that consumption and leisure are either complements or substitutes. The agent values bequest according to the utility function

$$u_d(\theta_d(t))=\frac{(k_d\theta_d(t))^{\alpha(1-\gamma)}}{1-\gamma} \tag{2.6}$$

with $k_d>0$. The constant k_d measures the intensity of preference for leaving a bequest. The parameters α,γ,k_d could be made dependent on the health state, but for ease of exposition we assume them constant in the baseline model.

Given initial wealth level $W(0)$ and health state $H(0)\in\{\mathrm{b},\mathrm{p}\}$, the agent's objective function is given by

$$\begin{aligned}\mathcal{J}(W(0),H(0);c,l,\boldsymbol{\theta},\tau_r)=E\Bigg[&\int_0^{\tau_d\wedge\tau_r}e^{-\delta t}u_c(c(t),l(t))\mathrm{d}t\\&+\int_{\tau_r\wedge\tau_d}^{\tau_d}e^{-\delta t}u_c(c(t),1)\mathrm{d}t\\&+e^{-\delta\tau_d}u_d(\theta_d(\tau_d-))\Bigg].\end{aligned}$$

By using the properties of the conditionally Poisson setting, the above can be rewritten as follows (e.g., Biagini, Biffis, Gozzi, & Zanella, 2022; Biffis, Gozzi, & Prosdocimi, 2020; Pham, 2009):

$$\begin{aligned}E\Bigg[&\int_0^{\tau_r}e^{-\int_0^t\beta(s)ds}u(c(t),l(t),\theta_d(t),H(t))\mathrm{d}t\\&+\int_{\tau_r}^{+\infty}e^{-\int_0^t\beta(s)ds}u(c(t),1,\theta_d(t),H(t))\mathrm{d}t\Bigg],\end{aligned} \tag{2.7}$$

where $\beta(s):=\delta+\lambda_d(H(s))$ denotes the mortality risk adjusted discount rate and where we have introduced the notation $u(c(t),l(t),\theta_d(t),H(t)):=u_c(c(t),l(t))+\lambda_d(H(t))u_d(\theta_d(t))$. As discussed in Biffis et al. (2020), Biagini et al. (2022), this shows that we can solve our optimization problem relative to the filtration $\mathbb{F}$. For ease of notation, we will keep on writing $(c,l,\boldsymbol{\theta},\tau_r)$

[3] This is assumed to be predictable relative to the information generated by the state variables.

Table 1
Mapping between variables in the primal and dual space.

Variable	Primal space	Dual Space
Retirement threshold	$\bar{x}$	$\bar{z}$
Threshold for leisure smaller than L	$\tilde{x}$	$\tilde{z}$
Optimal controls	$c^*, l^*, \theta_S^*, \theta_h^*, \theta_d^*$	z^*

for our controls, although what we will be using are their pre-death counterparts,[4] which are predictable relative to the smaller filtration $\mathbb{F}$.

We then define the agent's value function as

$$\mathcal{V}(W(0),H(0)) := \sup_{(c,l,\boldsymbol{\theta},\tau_r)\in\mathcal{A}} \mathcal{J}\big(W(0),H(0);c,l,\boldsymbol{\theta},\tau_r\big), \tag{2.8}$$

where $\mathcal{A}$ is the set of admissible strategies such that the problem is well-posed.

As in Farhi & Panageas (2007), Choi et al. (2008), and Dybvig & Liu (2010), it is convenient to write the objective (2.7) as follows:

$$E\left[\int_0^{\tau_r} e^{-\int_0^t \beta(s)ds} u(c(t),l(t),\theta_d(t),H(t))\,\mathrm{d}t \right. \\ \left. + e^{-\int_0^{\tau_r}\beta(s)ds}\, U(W(\tau_r),H(\tau_r))\right], \tag{2.9}$$

where $U(W(\tau_r),H(\tau_r))$ is the value function of the agent once she decides to retire:

$$U(W(\tau_r),H(\tau_r)) : = \sup_{(c,\boldsymbol{\theta})} E_{\tau_r}\left[\int_{\tau_r}^{+\infty} e^{-\int_{\tau_r}^t \beta(s)ds} \right. \\ \left. u(c(t),1,\theta_d(t),H(t))\mathrm{d}t\right]. \tag{2.10}$$

In line with Karatzas & Wang (2000), the idea is to solve first problems (2.9) and (2.10) separately for a fixed retirement date τ_r, wealth level $W(\tau_r)$, and health state $H(\tau_r)$, and then solve for the optimal retirement time τ_r^*. The next section outlines the solution strategy and illustrates the main features of the agent's optimal labor and portfolio choice.

3. Solution

In this section, we solve the problem by using martingale and duality methods. Use of the latter in portfolio choice problems originates from Pliska (1986), Karatzas, Lehoczky, & Shreve (1987), Cox & Huang (1989). The idea is to handle the dynamic budget constraint by introducing a semimartingale playing the role of a dynamic Lagrangian multiplier. By exploiting the martingale property of pricing functionals, optimization can be reduced to the static problem of finding the initial value of the Lagrangian semimartingale. A nice feature of the approach is that the Langrange multiplier can be interpreted as the shadow price of wealth as in static optimization problems. The case of incomplete markets (unspanned sources of risk and constraints) can be handled by pushing the approach further and using elegant convex duality techniques (see, for example, Pennanen, 2011, for an overview). As we will see in the following, the approach reduces the dimensionality of the problem by shrinking the set of controls $(c,l,\theta_S,\theta_h,\theta_d)$ to the starting value of the Lagrangian semimartingale, $z(0)$; Table 1 summarizes the mapping between primal and dual variables in our setting. We refer to Xu & Shreve (1992), He & Pages (1993), Karatzas & Shreve (1998), Choi et al. (2008), Dybvig & Liu (2010), Chen & Vellekoop (2017), Dong & Zheng (2020), Kamma & Pelsser (2022), among others, for applications to a variety of portfolio choice problems. Use of convex duality techniques seems to offer the only way to obtain semi-closed form solutions to our challenging problem, which is of mixed control and stopping type.

We now outline the solution strategy and the main results, while relegating detailed proofs to the online supplementary material. In line with He & Pages (1993), we introduce the state-price-density process ξ given by

$$\xi(t) = e^{-\left(r+\frac{1}{2}\Theta^2\right)t-\Theta Z(t)} e^{-\int_0^t \lambda_d(H(s))ds}, \tag{3.1}$$

where $\Theta := \frac{b-r}{\sigma}$ denotes the market price of financial risk,[5] and introduce the process $z(t) = \xi_0 e^{\int_0^t \beta(s)ds}\xi(t)$ having dynamics

$$\mathrm{d}z(t) = -(r-\delta)z(t)\mathrm{d}t - \Theta z(t)\mathrm{d}Z(t), \quad z(0) = \xi_0. \tag{3.2}$$

Here, the process z represents the dynamic Lagrange multiplier.

We then define the convex conjugate of the utility flow, u_c, and bequest function, u_d, as

$$\widetilde{u}_c(z,H(t)) := \max_{c\ge 0,\ 0\le l\le \widetilde{l}(t)} u_c(c,l) - (c + w(H(t))l^p)z, \tag{3.3}$$

where $\widetilde{l}(t) := \bar{l}(H(t))$ and

$$\widetilde{u}_d(z) := \max_{\theta_d} u_d(\theta_d) - \theta_d z, \tag{3.4}$$

so that we have $\widetilde{u}(z,H(t)) := \widetilde{u}_c(z,H(t)) + \lambda_d(H(t))\widetilde{u}_d(z)$. The processes $\widehat{c}, \widehat{l}, \widehat{\theta}_d$ solving problems (3.3)-(3.4) are given explicitly in the online supplementary material in Proposition A.4. Similarly, we define the convex conjugate of U as

$$\widetilde{U}(z,H(t)) := \sup_{\widehat{w}\ge 0} U(\widehat{w},H(t)) - \widehat{w}z. \tag{3.5}$$

By the optional sampling theorem, the budget constraint (2.4) for our optimization problem can be shown to take the following form (see section A.1 in the online supplementary material):

$$E\left[\int_0^{\tau_r} \xi(t)\Big\{c(t) - Y(w(H(t)),l(t)) + \lambda_d(t)\theta_d(t) \right. \\ \left. + e^{-\lambda_h t}\lambda_h(1-\eta(t))E[M]\Big\}\mathrm{d}t + W(\tau_r)\xi(\tau_r)\right] \le W(0). \tag{3.6}$$

Let us define

$$\widetilde{V}(W(0),H(0),\xi_0,\tau_r) : \\ = E\left[\int_0^{\tau_r} e^{-\int_0^t \beta(s)ds}\Big(\widetilde{u}(z(t),H(t)) \right. \\ + z(t)(w(H(t)) - (1-N_h(t))\lambda_h(1-\eta(t))E[M])\Big)\mathrm{d}t \\ \left. + e^{-\int_0^{\tau_r}\beta(s)ds}\widetilde{U}(z(\tau_r),H(\tau_r))\right]. \tag{3.7}$$

Then, for a fixed time τ_r, denote by $\mathcal{A}_{\tau_r}$ the set of admissible triplets $(c,l,\boldsymbol{\theta})$ for problem (2.8) and define the value function

$$V_{\tau_r}(W(0),H(0)) := \sup_{(c,\,l,\,\boldsymbol{\theta})\in\mathcal{A}_{\tau_r}} \mathcal{J}(W(0),H(0);c,l,\boldsymbol{\theta},\tau_r). \tag{3.8}$$

As demonstrated in Proposition A.2 in the online supplementary material, we can then prove that

$$V_{\tau_r}(W(0),H(0)) = \inf_{\xi_0>0}\Big[\widetilde{V}(W(0),H(0),\xi_0,\tau_r) + \xi_0 W(0)\Big]. \tag{3.9}$$

[4] In our framework the following result holds (see Aksamit & Jeanblanc, 2017, Proposition 2.11(b)): if a process A is $\mathbb{G}$-predictable then there exists a process a which is $\mathbb{F}$-predictable and such that $A(s,\omega) = a(s,\omega)$ for all $\omega\in\Omega$ and $s\in[0,\tau_d(\omega)]$. We refer to process a as to the *pre-death counterpart* of process A.

[5] In particular, under the equivalent pricing measure $\widetilde{\mathbb{P}}$, defined via $\widetilde{\mathbb{P}}(A) := E\big[\exp\big(-\Theta Z(t) - \frac{1}{2}\Theta^2 t\big)\,1_A\big]$, $A\in\mathcal{F}_t$, we have that $\widetilde{Z}(t) = Z(t) + \Theta t$ is a standard Brownian motion and the compensated jump processes $N_j(t) - \int_0^t(1-N_j(s))\lambda_j ds$ (for $j\in\{h,d\}$) are martingales. As our market setting is complete, the pricing measure is unique. Moreover, the doubly stochastic setting is preserved (see Biffis et al., 2010). See Section 5 for the extension to an incomplete market setting.

By (2.8) and (3.8), we can write

$$\mathcal{V}(W(0),H(0)) = \sup_{\tau_r\in[0,+\infty]} V_{\tau_r}(W(0),H(0)), \tag{3.10}$$

which, under the conditions of Proposition A.2, can be reduced to the static problem

$$\mathcal{V}(W(0),H(0)) = \sup_{\tau_r\in[0,+\infty]}\inf_{\xi_0>0}\Big[\widetilde{V}(W(0),H(0),\xi_0,\tau_r)+\xi_0 W(0)\Big] = \inf_{\xi_0>0}\Big[\overline{V}(\xi_0)+\xi_0 W(0)\Big], \tag{3.11}$$

where

$$\overline{V}(\xi_0) := \sup_{\tau_r\in[0,+\infty]} \widetilde{V}(W(0),H(0),\xi_0,\tau_r). \tag{3.12}$$

The value which gives the infimum in Eq. (3.11), ξ_0^*, is the optimal Lagrange multiplier associated to the static budget constraint.

To solve the optimal stopping problem (3.10), it is convenient to define

$$\begin{aligned}\phi_j(t,z) : = \sup_{\tau_r>t} E\Bigg[\int_t^{\tau_r} & e^{-\int_0^s\beta(u)\mathrm{d}u}\Big[\widetilde{u}(z(s),H(s)) \\ &+\Big(w(H(s))-\lambda_h E[M](1-\eta(s))\mathbf{1}_{H(s)=b}\Big)z(s)\Big]\mathrm{d}s \\ &+e^{-\int_0^{\tau_r}\beta(s)\mathrm{d}s}\widetilde{U}(z(\tau_r),H(\tau_r))\Big|z(t)=z,H(t)=j\Bigg],\end{aligned} \tag{3.13}$$

for $j\in\{\mathrm{b},\mathrm{p}\}$, and $\Phi_j(z) := e^{\int_0^t\beta(s)\mathrm{d}s}\phi_j(t,z)$, $j\in\{\mathrm{b},\mathrm{p}\}$, where the time-homogeneity of the Φ_j's follows from Barucci & Marazzina (2011). It holds that $\overline{V}(\xi_0)=\phi(0,\xi_0)$, with $\phi(t,z) := \sum_{j\in\{\mathrm{b},\mathrm{p}\}}\phi_j(t,z)1_{H(t)=j}$, and hence once the function ϕ is computed we can derive the optimal Lagrange multiplier as

$$\xi_0^* = \arg\min_{\xi_0>0}\phi(0,\xi_0)+\xi_0 W(0).$$

For ease of exposition, we now consider the simplification $\eta(t)=\eta_w 1_{\{\tau_r>t\}}+\eta_r 1_{\{\tau_r\le t\}}$, for deterministic parameters $\eta_w,\eta_r\ge 0$, which will be discussed in the analysis of Section 4.3. Following Barucci & Marazzina (2011), we can couple the results in He & Pages (1993) with those in Buffington & Elliott (2002) to show that the solution of the problem is obtained by solving a system of free-boundary problems in a regime-switching framework. In particular, we show that there exist boundaries $\bar{z}_\mathrm{b},\bar{z}_\mathrm{p}\in\mathbb{R}$ and functions $\Phi_\mathrm{p}\in C^1(\mathbb{R}_+)\cap C^2(\mathbb{R}_+\setminus\bar{z}_\mathrm{p})$, $\Phi_\mathrm{b}\in C^1(\mathbb{R}_+)\cap C^2(\mathbb{R}_+\setminus\bar{z}_\mathrm{b})$ (see also Choi et al., 2008, for lack of twice differentiability at the boundaries) such that

$$\begin{aligned}&\mathcal{L}_\mathrm{p}\Phi_\mathrm{p}(z)+\widetilde{u}_\mathrm{p}(z)+z\underline{w}=0 \quad \text{if } z>\bar{z}_\mathrm{p},\\ &\mathcal{L}_\mathrm{p}\Phi_\mathrm{p}(z)+\widetilde{u}_\mathrm{p}(z)+z\underline{w}\le 0 \quad \text{if } 0<z\le\bar{z}_\mathrm{p},\\ &\Phi_\mathrm{p}(z)\ge\widetilde{U}_\mathrm{p}(z) \quad \text{if } z>\bar{z}_\mathrm{p}, \quad \Phi_\mathrm{p}(z)=\widetilde{U}_\mathrm{p}(z) \quad \text{otherwise},\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}&\mathcal{L}_\mathrm{b}\Phi_\mathrm{b}(z)+\widetilde{u}_\mathrm{b}(z)+z(\overline{w}-\lambda_h(1-\eta_w)E[M])\\ &\qquad -\lambda_h\Phi_\mathrm{b}+\lambda_h\Phi_\mathrm{p}=0 \quad \text{if } z>\bar{z}_\mathrm{b},\\ &\mathcal{L}_\mathrm{b}\Phi_\mathrm{b}(z)+\widetilde{u}_\mathrm{b}(z)+z(\overline{w}-\lambda_h(1-\eta_w)E[M])\\ &\qquad -\lambda_h\Phi_\mathrm{b}+\lambda_h\Phi_\mathrm{p}\le 0 \quad \text{if } 0<z\le\bar{z}_\mathrm{b},\\ &\Phi_\mathrm{b}(z)\ge\widetilde{U}_\mathrm{b}(z) \quad \text{if } z>\bar{z}_\mathrm{b}, \quad \Phi_\mathrm{b}(z)=\widetilde{U}_\mathrm{b}(z) \quad \text{otherwise},\end{aligned} \tag{3.15}$$

where we recall that $\overline{w}$ and $\underline{w}$ are the wages defined in Section 2.2, and we have used the notation

$$\begin{aligned}&\mathcal{L}_j\Phi_j := -\beta_j\Phi_j+(\delta-r)z\frac{\partial\Phi_j}{\partial z}+\frac{1}{2}\Theta^2 z^2\frac{\partial^2\Phi_j}{\partial z^2},\\ &U_j(W(t)) := U(W(t),j), \qquad \widetilde{U}_j(z) := \widetilde{U}(z,j),\\ &u_j(c,l,\theta_d) := u(c,l,\theta_d,j), \qquad \widetilde{u}_j(z) := \widetilde{u}_c(z,j),\end{aligned} \tag{3.16}$$

for $j\in\{\mathrm{b},\mathrm{p}\}$. The solutions to variational inequalities (3.14)-(3.15) are given explicitly in Theorems A.8-A.9 in the online supplementary material.

We conclude with a verification theorem, for which we introduce the notation

$$\bar{z}(t) := \sum_{j\in\{\mathrm{b},\mathrm{p}\}}\bar{z}_j 1_{H(t)=j} \quad \text{and} \quad \Phi(t,z) := \sum_{j\in\{\mathrm{b},\mathrm{p}\}}\Phi_j(z)1_{H(t)=j}$$

(see Theorem A.3 in the online supplementary material). Here $z^*(t)$ is the process defined by the dynamic (3.2) with initial value ξ_0^*.

Theorem 3.1 (Verification Theorem). *Consider the pair $(\bar{z}(t),\Phi(t,z))\in\mathbb{R}_+\times C^1(\mathbb{R}_+)\cap C^2(\mathbb{R}_+\setminus\bar{z})$ and assume that it solves variational inequalities (3.14)-(3.15). Then, we can write $\phi_j(t,z)=e^{-\int_0^t\beta(s)ds}\Phi(t,z)$, and the latter coincides with $\phi_j(t,z)$ defined in (3.13), for $j\in\{\mathrm{b},\mathrm{p}\}$. Moreover, the optimal stopping time τ_r^* is given by*
$\tau_r^*=\inf\{s>t: z^*(s)\le\bar{z}(s)\}<+\infty$ a.s, where $z^*(t)=\xi_0^* e^{\int_0^t\beta(s)ds}\xi(t)$.

3.1. Optimal strategies

In this section, we present the optimal portfolio and insurance allocations, as well as labor market participation decisions.

Proposition 3.2 (Optimal labor supply). *There exist health dependent thresholds $\overline{W}_\mathrm{b}$ and $\overline{W}_\mathrm{p}$ such that the agent's optimal retirement time τ_r^* coincides with the first time when the optimal wealth path exceeds the wealth level $\overline{W}(t) := \overline{W}_\mathrm{b}1_{H(t)=\mathrm{b}}+\overline{W}_\mathrm{p}1_{H(t)=\mathrm{p}}$. At each time $t\ge 0$ before retirement, the optimal leisure is given by*

$$\begin{cases} l^*(t)=\left(\alpha z^*(t)^{-1}\left(\frac{\alpha p}{1-\alpha}w(H(t))\right)^{\alpha(1-\gamma)-1}\right)^{\frac{1}{\gamma+(1-p)(\alpha(1-\gamma)-1)}} & \text{if } z^*(t)\ge\tilde{z}(t),\\ l^*(t)=\bar{l}(H(t)) & \text{otherwise},\end{cases} \tag{3.17}$$

with the health-dependent boundary explicitly given by

$$\tilde{z}(t) := \alpha\left(\frac{\alpha p}{1-\alpha}w(H(t))\right)^{\alpha(1-\gamma)-1}\bar{l}(H(t))^{-\gamma-(1-p)(\alpha(1-\gamma)-1)}. \tag{3.18}$$

Proof. See section A.2 in the online supplementary material. □

Proposition 3.2 shows that the health state affects leisure and labor market participation in two ways. First, in line with Farhi & Panageas (2007), the distance to retirement is determined by the agent's current wealth relative to the retirement threshold $\overline{W}(t)$, which in our setting depends on the agent's health state. Second, before retirement labor supply is driven by the state variable z^*: if it is high enough, then labor supply is shaped by hourly wages and preference parameters; if it falls below a health dependent boundary, then the agent simply opts for maximal leisure. The better health state results in greater distance to retirement ($\overline{W}_\mathrm{b}>\overline{W}_\mathrm{p}$), as the agent needs to work longer to smooth consumption in the face of a longer life expectancy and medical expenses yet to be incurred. Similarly, before retirement the state variable z^* has greater slack in the better health state ($\tilde{z}_\mathrm{b}<\tilde{z}_\mathrm{p}$), thus increasing the distance to maximal leisure choice.

Similar trade-offs are at play when determining the optimal consumption strategy, which is given in the next proposition. If z^* is high enough, consumption is linear in hourly wages, and nonlinear in leisure choice. When z^* falls below the health dependent boundary $\tilde{z}$, then consumption is insensitive to wages and the agent simply opts for maximal leisure. Due to productivity losses (hourly wages drop to $\underline{w}$) and medical treatments (the maximal leisure available falls to $\underline{l}$), the worse the health status, the lower the consumption level afforded by the agent.

Proposition 3.3 (Optimal consumption)**.** *For each time $t \geq 0$ before the retirement time τ_r^*, the optimal consumption level is given by*

$$\begin{cases} c^*(t) = \frac{\alpha p}{1-\alpha} w(H(t))\, l^*(t)^p & \text{if } z^*(t) \geq \tilde{z}(t), \\ c^*(t) = \left(\alpha z^*(t) \bar{l}(H(t))^{(\alpha-1)(1-\gamma)}\right)^{\frac{1}{\alpha(1-\gamma)-1}} & \text{otherwise,} \end{cases} \tag{3.19}$$

where the health dependent boundary is given explicitly in (3.18).

Proof. See Proposition A.4 in section A.2 in the online supplementary material. □

From the above results we can see that retirement results in a drop in consumption as $l^*(t) \leq \bar{l}(H(t)) < 1$ if $t < \tau_r^*$ and leisure jumps to 1 after retirement; see Hubener, Maurer, & Mitchell (2015) for empirical evidence supporting this result. This downward jump ensures continuity of the marginal utility of consumption, a consequence of the smooth pasting principle (see Dybvig & Liu, 2010). As in Farhi & Panageas (2007), one can show that consumption prior to retirement is lower than in the case when no retirement option is available, as the agent needs to save more ahead of her exit from the labor market. Saving and insurance decisions are fully characterized in the next proposition.

Proposition 3.4 (Optimal investment, insurance, and bequest)**.** *The optimal investment in the risky asset is given by*

$$\theta_S^*(t) = \frac{\Theta}{\sigma} z^*(t) \frac{\partial^2}{\partial z^2} \Phi(t, z^*(t)), \tag{3.20}$$

for $t < \tau_r^*$.

The optimal health insurance demand θ_h^* is given by

$$\begin{aligned} \theta_h^*(t) = \frac{1}{\lambda_h E[M]} \Big\{ & (1 - l^*(t)^p) w(H(t)) - c^*(t) \\ & + (b - r - W^*(t)) \theta_S^*(t) \\ & - \lambda_d(H(t)) \left(\theta_d^*(t) - W^*(t)\right) + (\delta - r) z^*(t) \frac{\partial^2}{\partial z^2} \Phi(t, z^*(t)) \\ & + \frac{1}{2} (z^*(t)\Theta)^2 \frac{\partial^3}{\partial z^3} \Phi(t, z^*(t)) \Big\}, \end{aligned} \tag{3.21}$$

for $t < \min(\tau_r^*, \tau_h)$, and vanishes after the health shock occurrence.

Finally, the optimal bequest target θ_d^* is given by

$$\theta_d{}^*(t) = \left(\frac{1}{\alpha} z^*(t) k_d^{\alpha(\gamma-1)}\right)^{\frac{1}{\alpha(1-\gamma)-1}}. \tag{3.22}$$

Proof. See the online supplementary material. □

From (3.20), we see that the optimal risky asset allocation needs to take into account the current health state, as well as future possible state transitions. As shown in the following proposition, the solution of the post-retirement asset allocation problem is the same as that of the classical Merton problem with optimal bequest in case the health shock has already occurred (e.g., Dybvig & Liu, 2010). In case the health shock has not occurred yet, the solution takes into account medical expenses and the availability of health insurance. More precisely, the agent follows a Merton-type strategy in which financial wealth is replaced by total wealth (financial wealth plus human capital) net of the market value of future medical expenses; see Eq. (3.23) and section A.6 in the online supplementary material.

Proposition 3.5 (Post retirement strategies)**.** *From the retirement time τ_r^* onwards, the optimal consumption level is given by $c^*(t) = (\alpha z^*(t))^{\frac{1}{\alpha(1-\gamma)-1}}$, and the optimal investment in the risky asset and the optimal health insurance demand are given by Eqs. (3.20)-(3.21), respectively, replacing Φ with $\widetilde{U}$, where $\widetilde{U}$ admits the explicit health-state-dependent expressions:*

$$\widetilde{U}(z^*(t), \mathsf{p}) = \alpha^{\frac{1}{\Gamma}} \frac{\Gamma}{(1-\Gamma)} \frac{1 + \lambda_d(\mathsf{p}) k_d^{\frac{1-\Gamma}{\Gamma}}}{\xi_\mathsf{p}} z^*(t)^{\frac{\Gamma-1}{\Gamma}}$$

$$\widetilde{U}(z^*(t), \mathsf{b}) = \alpha^{\frac{1}{\Gamma}} \frac{\Gamma}{(1-\Gamma)} M_\mathsf{b}^{\frac{1}{\Gamma}} z^*(t)^{\frac{\Gamma-1}{\Gamma}} - \frac{\lambda_h}{r + \lambda_h} (1 - \eta_r) E[M] z^*(t),$$

with $\Gamma = 1 - \alpha(1-\gamma)$, $\xi_\mathsf{p} = \frac{\Gamma-1}{\Gamma}\left(r + \lambda_d(\mathsf{p}) + \frac{\Theta^2}{2\Gamma}\right) + \frac{\delta + \lambda_d(\mathsf{p})}{\Gamma}$, and constants M_b given in Section A.6 in the online supplementary material. Here, $\widetilde{U}$ are the convex conjugates of the post retirement utility functions, given by

$$U(W, \mathsf{p}) = \alpha \left(\frac{1 + \lambda_d(\mathsf{p}) k_d^{\frac{1-\Gamma}{\Gamma}}}{\xi_j} \right)^{\Gamma} \frac{W^{1-\Gamma}}{1-\Gamma},$$

and

$$U(W, \mathsf{b}) = \alpha M_\mathsf{b} \frac{\left(W - \frac{\lambda_h}{r+\lambda_h}(1-\eta_r) E[M]\right)^{1-\Gamma}}{1-\Gamma}. \tag{3.23}$$

Proof. See Dybvig & Liu (2010) and the online supplementary material. □

As pointed out in Karatzas & Wang (2000), a solution to the optimization problem (2.8) may not exist. The next theorem provides a simple sufficient condition for existence based on the agent's subjective discount rate being sufficiently high. Equivalently, the agent's mortality risk adjusted discount rate needs to be larger than a given threshold shaped by the preference parameters and the risk-return trade-off offered by the investment opportunities.

Theorem 3.6 (Sufficient condition for existence)**.** *Assume that the following condition is satisfied for all $t \geq 0$:*

$$\beta(t) > \beta^*, \tag{3.24}$$

where $\beta(t) = \delta + \lambda_d(H(t))$ and $\beta^* = \alpha(1-\gamma)\left(r + \lambda + \frac{1}{1-\alpha(1-\gamma)} \frac{\Theta^2}{2}\right)$. Then, the optimization problem (2.8) admits a solution.

Proof. See Section A.4 in the online supplementary material. □

A stronger condition implying (3.24) is $\delta + \lambda > \beta^*$. This is immediate to check and simply states that the mortality risk adjusted discount rate in the b health state should be greater than the threshold β^*.

4. Model estimation

We estimate the model by using the Health and Retirement Study (HRS), which covers a representative sample of older households surveyed every two years since 1992. The data source is perfectly suited for our model, as it follows households over time and provides information on health expenses, health outcomes, insurance holdings, as well as information on income and wealth. In line with French & Jones (2011), we focus our attention on optimal portfolio and labor market decisions made by an average 60-year-old male US individual throughout his working life, as well as after retirement. Conditioning on relevant socio-economic characteristics yields different calibrated parameters, but the main trade-offs supported by the optimal strategies discussed below remain unaltered. French & Jones (2011) build an index capturing preference heterogeneity in the sample population and repeat the results for three subclasses of agents obtaining similar results.

4.1. Estimation strategy

We adopt a two-stage estimation procedure. In the first stage, we estimate those parameters that can be identified without resorting to our model. These include mortality and health shock rates, as well as financial market parameters. In the second stage, we estimate the remaining parameters by using the simulated method of moments (e.g., Adda, Cooper, & Cooper, 2003, Chapter 4.3.3).

We estimate the components of the intensity of mortality by using a simple Markov chain model. The results, reported in Table 2, suggest that the health shock occurs on average after 9.29 years and that the life expectancy for a 60-year-old is 14.88 years.[6] Market parameters are estimated by using S&P500 and 3-month T-Bills monthly data during the period 1950–2006 for the risky and riskless asset, respectively. Labor income parameters $\overline{w}$ and $\underline{w}$ are estimated based on the `riearn` field of HRS data. In line with Aaronson & French (2004), we set $p = 2$, meaning that switching from full time to half time (i.e., increasing leisure from 0 to 50%) would generate a 25% drop in gross income. For simplicity, income is taxed at a flat 40% rate; see French & Jones (2011) for a more complex tax schedule. To estimate average medical expenses, we consider the cumulative costs[7] expected to be incurred over 10 years by a 60-year-old male individual without Medicare, discounted at the risk free rate. For this exercise, we therefore consider $\eta(t) \equiv 0$ in (2.4), i.e., no coverage of medical expenses from employer provided insurance or Medicare. The case with η different from zero is then considered in Section 4.3.

In the second stage, we estimate the remaining parameters $(k_d, \delta, \alpha, \gamma, \bar{l}, \underline{l})$ by using the method of simulated moments. In particular, we consider parameter values minimizing the distance between the mean values of three endogenous model outputs and the corresponding empirical values. The endogenous model outputs are the number of working hours, the optimal wealth level, and the optimal risky asset allocation.[8] For simplicity, the latter is obtained by considering the aggregate value of holdings in stocks, bonds, and pension savings. Cash and short term deposits are instead considered part of the riskless asset allocation. In line with French & Jones (2011), we consider 4060 hours as the annual leisure endowment, and define optimal working hours as given by $(1 - l^*(t)) \times 4060$. In carrying out our simulations, we allow for both Brownian and conditional Poisson randomness. The estimation error to be minimized is defined as a weighted sum of the L^2-norm calibration errors for expected wealth, leisure and risky asset allocation. When considering the average wealth level, for example, its calibration error is given by

$$Err_W = \frac{\sqrt{\sum_{i=0}^{10} |\overline{W}_{HRS}(60+2i) - \overline{W}_m(2i)|^2}}{\sqrt{\sum_{i=0}^{10} \overline{W}_{HRS}(60+2i)^2}},$$

where $\overline{W}_{HRS}$ and $\overline{W}_m$ correspond to the average wealth levels obtained from HRS data and our model, respectively. More precisely, $\overline{W}_m$ is computed on the basis of 20 000 simulations for the state variables, averaging among all the obtained wealth paths. We define the calibration error for leisure, Err_l, and for the stock market participation, Err_S, in a similar way. The errors are then weighted so as to maximize the goodness of fit across the three curves.

Table 2 reports the parameter estimates obtained across the two stages. The calibration error on the second stage is equal to 0.083. As an example, Fig. 1 depicts the matching result for the first moment of one of the three calibrated quantities, the labor market participation. To better capture the cross-sectional heterogeneity of households we could extend the simulated method of moments to target higher moments of the relevant variable's distribution.

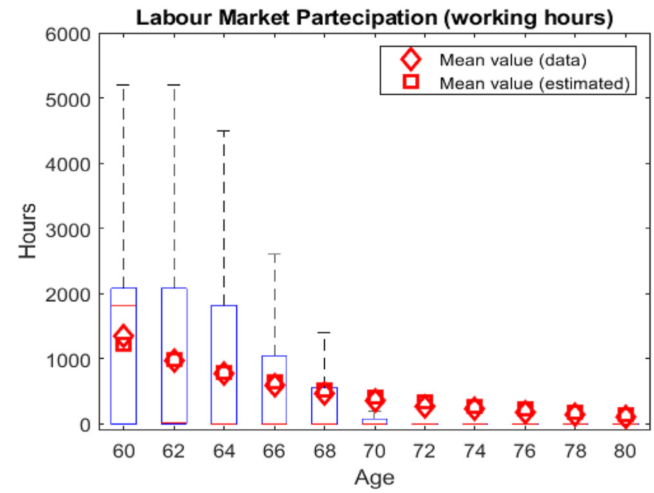

Fig. 1. HRS data boxplot and calibration results for the calibrated leisure, transformed in $(1 - l^*(t)) \times 4060$ working hours. Mean values from the sample, mean values are estimated on the basis of 20 000 simulations for the state variables.

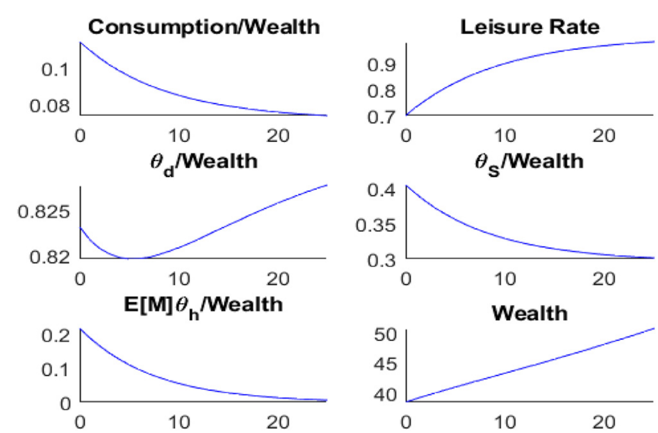

Fig. 2. Private insurance. Optimal insurance, investment, and consumption strategies (as a fraction of optimal wealth), as well as optimal leisure for a average agent aged 60 in the baseline model of Section 3. Average values are computed on the basis of 200 000 simulations for the state variables. x-axis: age from 0 (60 years old) to 25 (85 years old). Optimal wealth is in 10 000 USD.

4.2. Optimal strategies: Private health insurance

To better understand the optimal strategies delivered by the model, we first consider the case in which only private insurance is available, i.e., we set $\eta(t) \equiv 0$ in (2.4). For comparison, we then consider a version of the framework in which we remove health risk and medical expenses altogether ($\lambda_h = 0$). Figure 2 reports optimal strategies for an (average)[9] 60-year-old male agent; the same strategies are reported in Fig. 3 for the case of $\lambda_h = 0$. We note that in both cases the model induces a decreasing pattern in the risky asset allocation as the (average) agent approaches retirement, i.e., as more and more simulations deal with the retirement option. The result applies also to the case of working agents. The latter

[6] The life expectancy is computed on the basis of one million simulations of the conditionally Poisson process with intensity $\lambda_d(H(t))$.

[7] This is in line with the annual average expenses for a person in Bad Health reported in French & Jones (2011).

[8] The relevant HRS variables are `rjhours*rjweeks`, `hatota`, and `hastck+habond+haira`, respectively.

[9] In the following, we refer to "average" agents when considering results averaging across "working" and "retired" agents along the optimal paths.

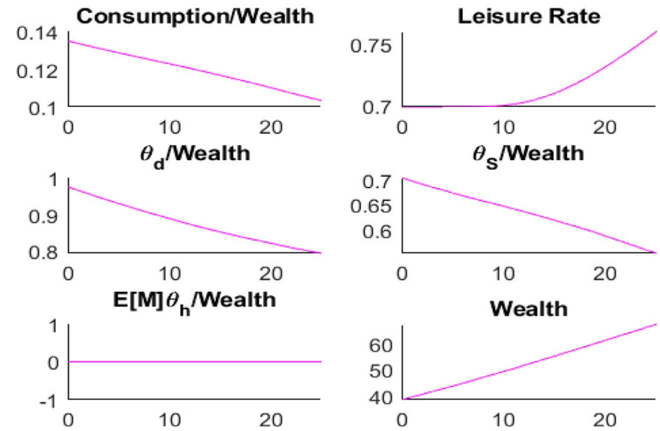

Fig. 3. No health risk. Optimal insurance, investment, and consumption strategies (as a fraction of optimal wealth), as well as optimal leisure for a average agent aged 60 in the model without health shock. Average values are computed on the basis of 200 000 simulations for the state variables. *x*-axis: age from 0 (60 years old) to 25 (85 years old). Optimal wealth is in 10 000 USD.

Table 2
Main parameter values. $\overline{w}$, $\underline{w}$, M and $W(0)$ are given in 10 000 USD.

Health Shock	$\lambda = 0.01376$, $\lambda_h = 0.10765$, $\Delta_h = 0.11971$
Financial Market	$r = 0.0492$, $b = 0.0872$, $\sigma = 0.1426$
Labor income	$p = 2$, $\overline{w} = 6.4188$, $\underline{w} = 5.4237$
Medical Expenses	M=8.7055
Preferences	$\delta = 0.0190$, $\alpha = 0.3880$, $\gamma = 14.6978$
Maximum Leisure	$\bar{l} = 0.6991$, $\underline{l} = 0.3046$
Bequest	$k_d = 0.0542$
Initial Wealth	$W(0) = 38.417$

Table 3
Retirement thresholds for an (average) agent aged 60 at time 0 and model parameters as in Table 2.

Health insurance	b	p
Private insurance ($\eta_w = \eta_r = 0$)	94.61	16.13
Tied health insurance ($\eta_w = 1, \eta_r = 0$)	119.27	16.13
Retiree health insurance ($\eta_w = \eta_r = 1$)	86.44	16.13
Medicare ($\eta_w = 0, \eta_r = 1$)	61.76	16.13

has traditionally been regarded as a puzzle (relative to the empirical predictions of Merton-like models) and then solved by looking, for example, at the co-integration between labor income and stock prices (see Benzoni, Collin-Dufresne, & Goldstein, 2007). Here, we see an alternative explanation in which the compression in risky asset holdings is induced by the agent gradually investing more conservatively to be able to retire.

When medical expenses are considered, health insurance demand is sizeable, thus decreasing the resources available for investment and leading to a more pronounced compression in risky asset investment. Moreover, in line with De Nardi, French, & Jones (2010), the agent consumes less and the fraction of wealth allocated to life insurance is smaller and non-linear, as bequest motives are traded off against preferences for leisure during retirement.

Table 3 reports the health-dependent retirement thresholds relative to different types of insurance, including the private health insurance case considered here. We find that an average 60-year-old male agent retires after 9.1 years, at age 69.1. There is a wedge between the thresholds in the case with health risk (state b) and without (state p, for which the health shock has already occurred). The sizeable wedge is mostly due to the simple (binary) nature of health risk assumed in our baseline model. The message is that the prospect of health shocks and associated medical expenses induces the agent to work longer, whereas the agent can retire immediately as soon as such expenses are no longer material (see Chai et al., 2011, for related results). The optimal strategies obtained demonstrate that working longer is not just the result of a higher retirement threshold, but also of the slower wealth accumulation along the optimal path due to health insurance purchases. Similarly, optimal consumption and labor market participation decrease in the presence of health risk, consistently with the empirical evidence documented in French & Jones (2011).

4.3. Optimal strategies: Employer-provided health insurance

We then extend the baseline model to include exogenous insurance provision, which is characterized by the process $\eta(t) \geq 0$ and will be seen to provide important retirement incentives. In line with French & Jones (2011), we consider two forms of employer-provided health coverage:

- *Tied* health insurance coverage, which is provided by the employer while the agent is actively working: $\eta(t) > 0$ on the event $\{\tau_r > t\}$ and $\eta(t) = 0$ on the event $\{\tau_r \leq t\}$.
- *Retiree* health insurance coverage, meaning that coverage is retained by the employee also after retirement: $\eta(t) > 0$.

For comparison we also consider the following cases:

- *Private health insurance only*. As in the baseline model of the previous section, we set $\eta(t) = 0$.
- *Medicare*. No employer-provided insurance, governmental health coverage after retirement: $\eta(t) = 0$ on $\{\tau_r > t\}$ and $\eta(t) > 0$ on $\{\tau_r \leq t\}$. We regard this case as a proxy[10] for Medicare in the US and other forms of governmental support elsewhere.

We assume each form of coverage to deliver a payout equal to $\eta(t)M$, conditional on the health shock occurring. We have full coverage in case $\eta(t) = 1$. For ease of illustration, in the following discussion we consider the simple parametrization $\eta(t) = \eta_w 1_{\{\tau_r > t\}} + \eta_r 1_{\{\tau_r \leq t\}}$, for deterministic parameters $\eta_w, \eta_r \geq 0$.

Table 3 reports the retirement wealth thresholds conditional on different health states. The retirement threshold for state p is constant, as health coverage becomes immaterial once the health shock already occurred. Retirement thresholds differ considerably in health state b instead. In line with the analysis provided in Rust & Phelan (1997) and French & Jones (2011), we see that agents set for themselves a higher wealth threshold for retirement in the *tied* coverage case relative to the case of *private* health insurance only. Agents target lower wealth thresholds for retirement both in the *retiree* and *Medicare* case, the latter providing stronger work disincentives. To properly understand how wealth thresholds translate into retirement decisions, we need to consider the relevant optimal strategies and associated optimal wealth paths, which are reported in Fig. 4. We note that the optimal strategies for the *Medicare* and *private* coverage cases are very close. This is largely due to the strategy needing to support private health insurance demand at the expense of wealth, consumption, risky asset allocation, and bequest before retirement. Despite the similarity of the optimal strategies, agents retire earlier on average in the *Medicare* case (lower optimal wealth threshold) as they can gain access to coverage during retirement. Considering now the cases of *tied* and *retiree* health insurance coverage, they also present similar optimal strategies: lower health insurance demand, higher wealth, consumption, bequest and risky asset allocation compared to the *private* and *Medicare* case. The *tied* case results in later retirement (higher wealth threshold), as agents reap the benefits of

[10] We make the simplification of not considering age-contingent eligibility requirements, currently set at age 65 for Medicare, for example.

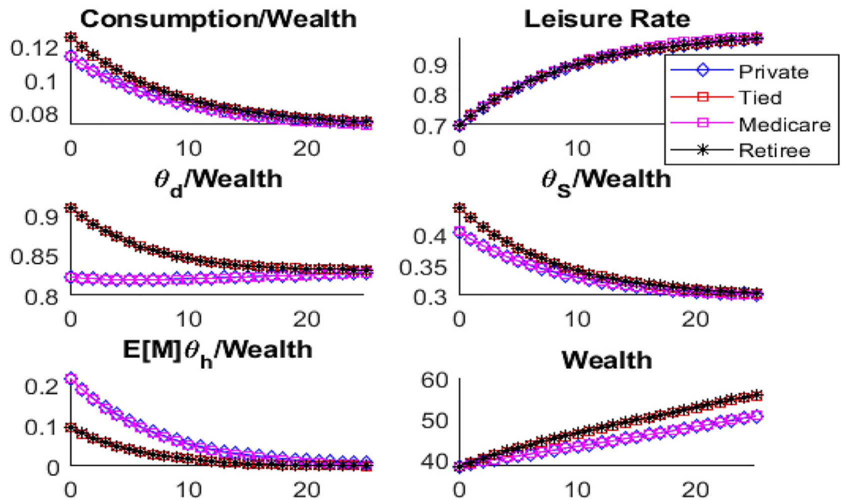

Fig. 4. Optimal strategies for an (average) agent aged 60 (see Table 2) with initial health state b for different types of insurance coverage. Average values are computed on the basis of 200 000 simulations for the state variables. *x*-axis: age from 0 (60 years old) to 25 (85 years old). *y*-axis: all the strategies (with the exception of the leisure rate) are in 10 000 USD.

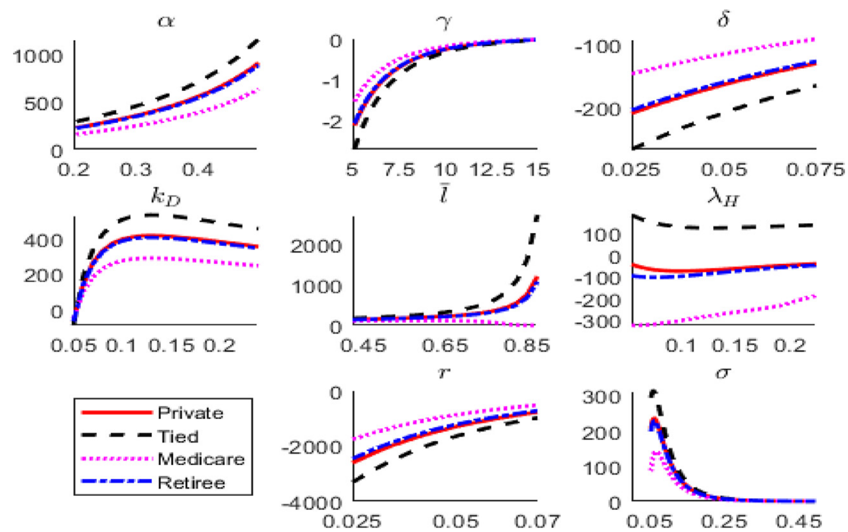

Fig. 5. First order derivative of the retirement threshold in the b state relative to different parameters, all other parameters being set as in Table 2.

employer-provided health insurance for as long as they manage to build a large enough buffer to support medical expenses and consumption during retirement. *Retiree* coverage is associated with a much lower optimal wealth threshold for retirement, as medical expenses are covered by the employer.

4.4. Sensitivity analysis

We now explore the sensitivity of optimal retirement thresholds to changes in key model parameters. In Fig. 5, we report the first order derivative of the wealth threshold (delta) relative to the preference parameters (α, γ, δ), bequest intensity (κ_D), leisure (l_b), health shock intensity (λ_H), and market parameters (r, σ). We remark that any result pertaining to retirement thresholds must be understood in relation to the corresponding optimal wealth path.

The results show sizeable non-linear effects at play for several parameters. We discuss a selected few:

- Interest rate level. It is negatively related to the retirement threshold. A higher riskless rate environment allows wealth to accumulate faster, but also makes future medical expenses and consumption streams smaller in present value terms. The agent can therefore retire earlier.
- Bequest intensity. The agent works longer to accumulate greater resources to leave as bequest, hence the bequest intensity is positively related to labor supply (higher retirement thresholds).
- Subjective discount rate. The parameter δ is negatively related to the retirement thresholds: the agent retires earlier when the present value of future consumption and expenses are reduced because of a high discount rate.
- The probability of health shock occurrence. The impact on labor supply and retirement decisions can only be understood by considering the relevant health insurance framework. For example, in the case of *tied* coverage an increase in the likelihood of health shock occurrence (higher λ_H) makes expected medical expenses more material and therefore the agent delays retirement to extract the health insurance benefits provided by the employer. The opposite is true for the case of *Medicare*, as availability of health insurance during retirement provides an incentive to retire earlier. We note that the cases of *private* and *retiree* coverage

Table 4
Health states and mortality intensity levels as a function of spanned and unspanned health shock occurrences. The health states are: "best" on the event $\{N_a(t) = N_h(t) = 0\}$, "good" on $\{N_a(t) = 1, N_h(t) = 0\}$, "poor" on $\{N_a(t) = 0, N_h(t) = 1\}$, and "worst" on $\{N_a(t) = N_h(t) = 1\}$. In the baseline model, irreversible transitions can occur between states b and p, and between states b, g, and w.

	On $\{N_h = 0\}$	On $\{N_h = 1\}$
On $\{N_a = 0\}$	$H(t) = \mathsf{b}$ $\lambda_d(\mathsf{b}) = \lambda$	$H(t) = \mathsf{p}$ $\lambda_d(\mathsf{p}) = \lambda + \Delta_h$
On $\{N_a = 1\}$	$H(t) = \mathsf{g}$ $\lambda_d(\mathsf{g}) = \lambda + \Delta_a$	$H(t) = \mathsf{w}$ $\lambda_d(\mathsf{w}) = \lambda + \Delta_a + \Delta_h$

make the retirement threshold relatively insensitive to the health shock probability. The reason is that in both cases health coverage is provided through the agent's lifetime. As private insurance is fairly priced in our baseline model, there is no major difference between the two situations, although one must recall that the distance to retirement in the *private* coverage case is considerably larger than in the *retiree* case, as private insurance purchases make wealth accumulation slower.

5. Extension: Unspanned health risk

As an extension of the baseline model, we consider the possibility of including an unspanned source of health risk. Beyond standard life and health insurance repricing risk considerations, what we would like to capture with this model extension is health deterioration which is not associated with immediate medical expenses, but with a decrease in life expectancy and an increase in expected future medical expenses. We regard such health deterioration as a by-product of the natural "**a**ging" process undergone by any individual and therefore label any associated quantities with the subscript "a". The simplest way to capture this extra risk dimension is by extending the model to four possible health states, which are now labelled as follows: b for "best"; g for "good"; p for "poor"; w for "worst". An agent in the best health state, $H(t) = \mathsf{b}$, has intensity of mortality $\lambda > 0$. At an independent Poisson time τ_a with parameter $\lambda_a > 0$, the agent's health switches from state b to state g and the death intensity increases by $\Delta_a \geq 0$. A health shock can still occur at an independent Poisson time τ_h with parameter $\lambda_h > 0$, leading to a jump Δ_h in the current mortality intensity. For simplicity and ease of exposition, we assume here that the unspanned health shock can only occur if τ_h has not occurred yet. The resulting health states are $H(\tau_h) = \mathsf{p}$ if $H(\tau_h-) = \mathsf{b}$ and $H(\tau_h) = \mathsf{w}$ if $H(\tau_h-) = \mathsf{g}$. See Table 4 for a summary. The agent's intensity of mortality can now be written as follows:

$$\lambda_d(t) := \lambda + \Delta_a N_a(t) + \Delta_h N_h(t), \tag{5.1}$$

where $N_h(t)$ and $N_a(t)$ denote the health and natural aging shock indicator processes, respectively, the latter being defined for simplicity by $N_a(t) := 1_{\tau_a \leq t} 1_{\tau_h > \tau_a}$. In line with empirical evidence showing that medical expenses are higher on average for old age individuals (e.g., De Nardi et al., 2010), we make medical expenses depend on the health state by using the simple representation

$$M(t) = \overline{M} N_a(t) + \underline{M}(1 - N_a(t)),$$

with random variables $\overline{M}$ and $\underline{M}$ satisfying $E[\overline{M}] \geq E[\underline{M}]$ and representing random expenses of different average magnitude.

The introduction of an unspanned source of risk certainly makes the model more realistic but also introduces significant technical challenges as the market is no longer complete. In particular, there is no tradable instrument allowing one to hedge the jump Δ_a; the agent is therefore exposed to life and health insurance repricing risk. We solve the problem by adopting the fictitious market completion approach of Karatzas, Lehoczky, Shreve, & Xu (1991). In particular, we introduce a fictitious insurance product allowing the agent to hedge the aging shock via a wealth allocation denoted by $\theta_a(t)$. By the latter, we mean the payment at time t of a premium of amount $\theta_a(t)\lambda_a$ to allow the agent to receive a payout $\theta_a(t)$ in case the unspanned health shock occurs over the next small time interval. Among all the candidate pricing measures equivalent to $\mathbb{P}$ (equivalent martingale measures), we can next choose the one making the allocation to the fictitious hedge vanish, i.e., $\theta_a^*(t) = 0$. Intuitively, this ensures that the candidate strategy is the optimal one for the incomplete market model. See Remark A.5 in section A.3 in the online supplementary material for details.

Table 5
Main parameter values for the extended model. $\overline{w}$, $\underline{w}$, $\overline{M}$, $\underline{M}$ and $W(0)$ are given in 10 000 USD.

Aging Process*	$\lambda = 0.01376$, $\lambda_a = 0.200$, $\Delta_a = 0.03175$
Health Shock	$\lambda_h = 0.10765$, $\Delta_h = 0.11971$
Financial Market	$r = 0.0492$, $b = 0.0872$, $\sigma = 0.1426$
Labor income	$p = 2$, $\overline{w} = 6.4188$, $\underline{w} = 5.4237$
Medical Expenses*	$\underline{M} = 7.5700$, $\overline{M}$=9.8410
Preferences*	$\delta = 0.0267$, $\alpha = 0.3986$, $\gamma = 12.2632$
Maximum Leisure*	$\bar{l} = 0.7229$, $\underline{l} = 0.2030$
Bequest*	$k_d = 0.0426$
Initial Wealth	$W(0) = 38.417$

Table 6
Optimal retirement thresholds for an average top-wealth-quartile agent aged 60 and model parameters as in Table 5.

Health insurance	b	g	p	w
Private insurance	102.5740	104.9411	11.3040	11.0291
Tied insurance	129.8759	135.7914	11.3040	11.0291
Retiree insurance	91.6448	89.2351	11.3040	11.0291
Medicare	66.0641	58.4098	11.3040	11.0291

We calibrate the extended model to the 'average' 60-year old agent considered in the previous section. The relevant parameters are reported in Table 5.[11] The estimation error of the simulated method of moments procedure is comparable to the one obtained for the baseline model (with a relative error equal to 0.076 against a value of 0.083 for the baseline model).

The optimal retirement wealth thresholds for the extended model are reported in Table 6. An analysis of the results across the different health states reveals a divergent role of unspanned and spanned health shocks in shaping labor market participation. In general, the unspanned health shock considered here results in a lower life expectancy, thus allowing the agents to revise downwards their distance to retirement. However, the increase in expected future medical expenses means that agents will value labor income more as a way to support future healthcare costs. Our results suggest that the latter effect dominates and increases the distance to retirement in the *private* and *tied* insurance cases. The effect is the opposite in the *Medicare* and *retiree* cases, as the unspanned health shock accelerates the retirement decision in view of the support toward healthcare costs that will be provided after retirement.

6. Conclusion

In this paper we have developed a continuous time model shedding light on how medical expenses and health insurance shape optimal portfolio choice, labor supply, and retirement decisions. The model is simple enough to deliver solutions available in semi-analytic form, yet rich enough to match some important em-

[11] We set $\overline{M}$ and $\underline{M}$ such that $(\overline{M} + \underline{M})/2 = M$, and $M = 1.15\underline{M}$.

pirical evidence on portfolio decisions and labor market participation during the last phase of an agent's working life. The tractability of the model allows us to better understand incentives and disincentives provided by different health insurance schemes, ranging from private health insurance to employer-provided insurance and government schemes such as Medicare. The results demonstrate how ill designed policy interventions can result in undesirable fiscal effects resulting from the ability of workers to adjust labor supply. Although the findings are robust to several extensions of the model, including market incompleteness, it would be interesting to consider the possibility of (partially) endogenizing health dynamics as a result of health investment (e.g., Hugonnier et al., 2013). Studying the implications of this aspect for optimal retirement decisions is left for future research.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2022.09.016.

References

Aaronson, D., & French, E. (2004). The effect of part-time work on wages: Evidence from the social security rules. *Journal of Labor Economics, 22*(2) 329–252.
Adda, J., Cooper, R., & Cooper, R. W. (2003). *Dynamic economics: Quantitative methods and applications*. MIT press.
Aksamit, A., & Jeanblanc, M. (2017). *Enlargement of filtration with finance in view*. Springer.
Asmussen, S., & Steffensen, M. (2020). *Risk and insurance*. Springer.
Barucci, E., & Marazzina, D. (2011). Optimal investment, stochastic labor income and retirement. *Applied Mathematics and Computation, 218*(9), 5588–5604.
Bensoussan, A., Jang, B.-G., & Park, S. (2016). Unemployment risks and optimal retirement in an incomplete market. *Operations Research, 64*(4), 1015–1032.
Benzoni, L., Collin-Dufresne, P., & Goldstein, R. S. (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *The Journal of Finance, 62*(5), 2123–2167.
Biagini, S., Biffis, E., Gozzi, F., & Zanella, M. (2022). Wage rigidity and retirement in portfolio choice. *Working paper*.
Biffis, E., Denuit, M., & Devolder, P. (2010). Stochastic mortality under measure changes. *Scandinavian Actuarial Journal, 2010*(4), 284–311.
Biffis, E., Gozzi, F., & Prosdocimi, C. (2020). Optimal portfolio choice with path dependent labor income: The infinite horizon case. *SIAM Journal on Control and Optimization, 58*(4), 1906–1938.
Blau, D., & Gilleskie, D. (2008). The role of retiree health insurance in the employment behavior of older men. *International Economic Review, 49*(2), 475–514.
Blundell, R., French, E., & Tetlow, G. (2016). Chapter 8 - retirement incentives and labor supply. In *Handbook of the Economics of Population Aging: vol. 1* (pp. 457–566). North-Holland.
Bodie, Z., Merton, R. C., & Samuelson, W. F. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control, 16*(3), 427–449.
Buffington, J., & Elliott, R. J. (2002). American options with regime switching. *International Journal of Theoretical and Applied Finance, 5*(05), 497–514.
Chai, J., Horneff, W., Maurer, R., & Mitchell, O. S. (2011). Optimal portfolio choice over the life cycle with flexible work, endogenous retirement, and lifetime payouts*. *Review of Finance, 15*(4), 875–907. https://doi.org/10.1093/rof/rfr016.
Chen, A., & Vellekoop, M. (2017). Optimal investment and consumption when allowing terminal debt. *European Journal of Operational Research, 258*(1), 385–397.
Chen, C.-C., Chang, C.-C., Sun, E. W., & Yu, M.-T. (2022). Optimal decision of dynamic wealth allocation with life insurance for mitigating health risk under market incompleteness. *European Journal of Operational Research, 300*(2), 727–742. https://doi.org/10.1016/j.ejor.2021.10.016.
Choi, K. J., Shim, G., & Shin, Y. H. (2008). Optimal portfolio, consumption-leisure and retirement choice problem with CES utility. *Mathematical Finance, 18*, 445–472.
Cox, J. C., & Huang, C.-f. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory, 49*(1), 33–83. https://doi.org/10.1016/0022-0531(89)90067-7.
De Nardi, M. C., French, E., & Jones, J. B. (2009). Life expectancy and old age savings. *American Economic Review Papers and Proceedings*, 110–115.
De Nardi, M. C., French, E., & Jones, J. B. (2010). Why do the elderly save? The role of medical expenses. *Journal of Political Economy, 118*, 39–75.
Dong, Y., & Zheng, H. (2020). Optimal investment with s-shaped utility and trading and value at risk constraints: An application to defined contribution pension plan. *European Journal of Operational Research, 281*(2), 341–356.
Dybvig, P. H., & Liu, H. (2010). Lifetime consumption and investment: Retirement and constrained borrowing. *Journal of Economic Theory, 145*(3), 885–907.
Dynan, K. E., Skinner, J., & Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy, 112*, 397–444.
Farhi, E., & Panageas, S. (2007). Saving and investing for early retirement: A theoretical analysis. *Journal of Financial Economics, 83*(1), 87–121.
French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies, 72*(2), 395–427.
French, E., & Jones, J. B. (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica: Journal of the Econometric Society, 79*(3), 693–732.
Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy, 80*(2), 223–255.
Gustman, A., & Steinmeier, T. (1994). Employer-provided health insurance and retirement behavior. *Industrial and Labor Relations Review, 48*, 124–140.
Hausman, J. A., & Wise, D. A. (2008). Social security, health status, and retirement. In *Pensions, labor, and individual choice* (pp. 159–192). University of Chicago Press.
He, H., & Pages, H. F. (1993). Labor income, borrowing constraints, and equilibrium asset prices. *Economic Theory, 3*(4), 663–696.
Hoem, J. M. (1969). Markov chain models in life insurance. *Blätter der DGVFM, 9*(2), 91–107.
Hubbard, G., Skinner, J., & Zeldes, S. (1995). Precautionary saving and social insurance. *Journal of Political Economy, 103*(2), 360–399.
Hubener, A., Maurer, R., & Mitchell, O. S. (2015). How family status and social security claiming options shape optimal life cycle portfolios. *The Review of Financial Studies, 29*(4), 937–978.
Hugonnier, J., Pelgrin, F., & St-Amour, P. (2013). Health and (other) asset holdings. *Review of Economic Studies, 80*(2), 663–710.
Kamma, T., & Pelsser, A. (2022). Near-optimal asset allocation in financial markets with trading constraints. *European Journal of Operational Research, 297*(2), 766–781. https://doi.org/10.1016/j.ejor.2021.06.029.
Karatzas, I., Lehoczky, J. P., & Shreve, S. E. (1987). Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. *SIAM Journal on Control and Optimization, 25*(6), 1557–1586. https://doi.org/10.1137/0325086.
Karatzas, I., Lehoczky, J. P., Shreve, S. E., & Xu, G.-L. (1991). Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal on Control and optimization, 29*(3), 702–730.
Karatzas, I., & Shreve, S. E. (1998). *Methods of mathematical finance*: vol. 39. Springer.
Karatzas, I., & Wang, H. (2000). Utility maximization with discretionary stopping. *SIAM Journal on Control and Optimization, 39*(1), 306–329.
Koijen, R., Van Nieuwerburgh, S., & Yogo, M. (2016). Health and mortality delta: Assessing the welfare cost of household insurance choice. *The Journal of Finance, 71*(2), 957–1010.
Konicz, A. K., & Mulvey, J. M. (2015). Optimal savings management for individuals with defined contribution pension plans. *European Journal of Operational Research, 243*(1), 233–247.
Koo, B., Pantelous, A. A., & Wang, Y. (2022). Novel utility-based life cycle models to optimise income in retirement. *European Journal of Operational Research, 299*(1), 346–361.
Lee, L.-F. (1982). Health and wage: A simultaneous equation model with multiple discrete indicators. *International Economic Review, 23*(1), 199–221.
Lumsdaine, R. L., Stock, J. H., & Wise, D. A. (1996). Retirement incentives: The interaction between employer-provided pensions, social security, and retiree health benefits. In *The economic effects of aging in the united states and japan* (pp. 261–293). University of Chicago Press.
Maurer, R., Mitchell, O. S., Warshawsky, M. J., & Warshawsky, M. (2012). *Reshaping retirement security: Lessons from the global financial crisis*. Oxford University Press on Demand.
Merton, R. (1971). Optimal consumption and portfolio rules in a continuous time model. *Journal of Economic Theory, 3*, 373–413.
Owadally, I., Jang, C., & Clare, A. (2021). Optimal investment for a retirement plan with deferred annuities allowing for inflation and labour income risk. *European Journal of Operational Research, 295*(3), 1132–1146.
Pennanen, T. (2011). Convex duality in stochastic optimization and mathematical finance. *Mathematics of Operations Research, 36*(2), 340–362.
Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications*: vol. 61. Springer Science & Business Media.
Pliska, S. R. (1986). A stochastic calculus model of continuous trading: Optimal portfolios. *Mathematics of Operations Research, 11*(2), 371–382. https://doi.org/10.1287/moor.11.2.371.
Protter, P. E. (2005). *Stochastic differential equations*. Springer.
Rust, J., & Phelan, C. (1997). How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica: Journal of the Econometric Society*, 781–831.
Samuelson, P. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 239–246.
Xu, G.-L., & Shreve, S. (1992). A duality method for optimal consumption and investment under short- selling prohibition. I. General market coefficients. *The Annals of Applied Probability, 2*(1), 87–112.