The Geneva Risk and Insurance Review
https://doi.org/10.1057/s10713-024-00107-7

ORIGINAL ARTICLE

Market equilibrium with management costs and implications for insurance accounting

Michael Florig[1] · **Olivier Gossner**[2]

Received: 1 October 2022 / Accepted: 16 October 2024

Abstract

We examine a general equilibrium investment model in which agents incur management costs for holding assets. We characterize the influence of these costs on equilibrium prices as a weighted average of these costs for market participants. We then propose a correction method for this influence in valuation procedures used under regulatory frameworks, such as Solvency II. For insurers subject to Solvency II, the accounting correction amounts to approximately €130 billion, the equivalent of 1.8% of investments or 14% of own funds. These results not only contribute to the understanding of management costs in market equilibrium, but also highlight a distortion in current practices which discourages the holding of assets that are expensive to manage and typically inaccessible directly by policyholders.

Keywords General equilibrium · Insurance · Solvency II · Management costs · Valuation

JEL Classification D53 · G22 · L51 · M41

✉ Michael Florig
michael@florig.net

Olivier Gossner
olivier.gossner@polytechnique.edu

[1] Economics Department, École Polytechnique and Crédit Agricole CIB, Palaiseau, France

[2] CNRS-CREST, École Polytechnique and London School of Economics, Palaiseau, France

Published online: 13 December 2024

1 Introduction

Representing a vast segment of the financial sector, the insurance industry oversees a substantial asset base, chiefly comprising investments tied to long-duration retirement savings. Insurers within the European Economic Area (EEA) who are governed by Solvency II administer investments amounting to approximately €6,750 billion.[1] Insurers command a significant share of private household wealth. According to FFA (2021) and GDV (2020), it is estimated that French and German households, respectively, maintain 38% and 22% of their financial wealth through insurance contracts. Given the substantial volume of investments under their management, insurers serve as vital players in the distribution of capital across the economy.

Recent years have witnessed a global shift in insurance regulation toward a risk-based capital approach, underpinned by valuation methodologies developed within the academic literature, particularly arbitrage-free and market-consistent pricing. The Swiss Solvency Test (SST), instituted by the Swiss regulator in 2008, was an early adoption of this strategy, followed by the European Union implementing Solvency II throughout the European Economic Area (EEA) in 2016. Bermuda, an international insurance hub, implemented the Bermuda Solvency Capital Requirement (BSCR), deemed equivalent to Solvency II.

The valuation of market and investment-related risks in practice is achieved via (risk-neutral) pricing probabilities, first introduced in a general equilibrium setting by Drèze (1970) and further expanded notably by Ross (1977) and Cox and Ross (1976). Regulatory stipulations explicitly mandate the inclusion of investment management expenses in cash flow valuations.[2] These expenses not only contain directly incurred costs, like employee salaries and infrastructure costs, such as office space and software, but also external costs like fees paid to external fund managers and custody fees, typically subtracted from a fund's net asset value (EIOPA 2021).

A prevalent, though not universally accepted, method for meeting this regulatory requirement entails the valuation of investment-related cash flows under the implicit assumption of a financial market devoid of investment management costs, followed by the deduction of the present value of the company's expected future management costs under the same pricing measure. This approach has been documented as a common practice in CFO-Forum (2009)[3] for calculating shareholder value in a life insurance portfolio under the market-consistent embedded value approach, introduced by a consortium of European insurance CFOs. However, the management costs of market participants other than the company are not considered, resulting in an internal

[1] As per the Eiopa insurance statistics from 2024Q1. Additionally, these insurers manage approximately €2,300 billion of unit-linked investments on behalf of policyholders.

[2] See, e.g., EU Parliament (2015) §31.1: "A cash flow projection used to calculate best estimates shall take into account [⋯] investment management expenses".

[3] Compare to page 27, paragraph 137 of principle 13. Embedded value balance sheets were published by many listed European insurers before the introduction of Solvency 2. The embedded value should give a measure of the value from a shareholder perspective of a life insurance balance sheet, which mainly consists of a risk adjusted computation of discounted future cash flows paid to the shareholder. For complex life insurance products, this is typically based on risk-neutral Monte Carlo simulation techniques, similar to the valuation techniques used for exotic derivatives.

inconsistency in the Solvency II models typically implemented. Our analysis reveals that the standard method often leads to extensive double counting of costs.

The correction for the double counting in the current valuation methodology would be straightforward if all investors incurred identical management costs. The solution would simply involve ceasing to include the present value of individual management costs in the liability side, as these are already encapsulated in market prices. However, given that these costs vary across firms and regulations require the incorporation of individual management costs, a theory of financial markets and valuation that accommodates heterogeneous management costs is required to devise a consistent valuation approach. Although there is an important academic literature on transaction costs (see, e.g., Jouini and Kallal 1995; Cvitanić et al. 1999; Czichowsky et al. 2018), investment management costs have not received the same attention.

In this paper, we examine a financial economy where investors incur investment management costs, which can differ from one investor to another. As the market-consistent valuation of a portfolio hinges on observed market prices, our initial focus is to investigate the effect of management costs on equilibrium prices. We demonstrate that, in comparison to a cost-free scenario, the existence of costs deflates equilibrium prices by a factor measured by a weighted average of market participants' costs. The weights assigned to each market participant incorporate both their asset demand elasticities and the magnitude of their market position. Consequently, we develop a valuation formula for cash flows that takes into account the cost structure of the portfolio-holding company. This formula subtracts the company's individual costs from the market value of the cash flow-generating assets and reintroduces the weighted average of the market's management costs as a correction term.

The correction term is particularly relevant for (a) life insurance providers with investments underpinning long-term liabilities, and (b) insurers investing in assets that are complex to manage, with the impact being even more pronounced for insurers where both conditions apply. In fact, one of the significant potential value propositions of life insurers lies in their ability to offer access to such complex assets through a pooled investment process. The prevailing valuation approach, however, may not only overstate the effect of investment management costs on own funds but could also lead to unintentional repercussions, such as skewing investment strategies by incentivizing insurers to favor assets that are relatively inexpensive to manage.

The remaining structure of the paper is as follows: Sect. 2 introduces the financial model. Section 3 analyzes the sensitivity of equilibrium prices with respect to management costs. Section 4 applies our model to the valuation approach commonly adopted in the regulatory valuation practice and estimates the potential impact of our valuation approach. The paper concludes with Sect. 5.

2 Market equilibrium with management costs

We study a two-period exchange economy with a safe asset and a risky asset, similar to the models outlined in Fishburn and Porter (1976), Dana (1995), and Gollier (2001). Distinctively from previous models in the literature, our approach incorporates management costs associated with the risky investment. We characterize individual demand for the risky asset as a function of its price and management costs. By considering an arbitrary number of agents, we then establish the existence of a market equilibrium price under the general assumption of heterogeneous management costs among agents.

The finite set of agents is denoted I, the time periods are $t = 0, 1$, the economy has a safe asset, paying interest rate $\rho > -1$ between period 0 and 1, and a risky asset that pays a random amount x at period 1. These assets are traded at $t = 0$, and payoffs are realized at $t = 1$.

2.1 Individual demand

Each agent i holds $w_{ic} \geq 0$ in the riskless asset and $w_{ia} \geq 0$ in the risky asset at period 0, with at least one being positive. Agent i incurs a management cost of $m_i \geq 0$ for each unit of the risky asset, so their net payoff of the risky asset is $x - m_i$.

We assume the risky asset has bounded payoffs and does not lead to ruin: there exists a lower and upper bound $x_{\min}, x_{\max} > 0$ such that $m_i < x_{\min} \leq x \leq x_{\max}$ a.s.

We denote by U_i agent i's von Neumann–Morgenstern's utility function in numéraire, where

$$U_i : \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$$

with $u_i(w) \in \mathbb{R}$ for all $w > 0$.

For each i, U_i is twice differentiable, $U_i' > 0$ and $U_i'' < 0$. Agent i's Arrow–Pratt coefficient of relative risk aversion is $R_i^r(w) = -wU_i''(w)/U_i'(w)$. The agent has constant absolute risk aversion (CARA) preferences, if the coefficient of absolute risk aversion $R_i^a(w) = -U_i''(w)/U_i'(w)$ is independent of w and decreasing absolute risk aversion (DARA) preferences, if $R_i^a(w)$ is non-increasing in w.

Given a unit price $p > 0$ and a quantity a in the risky asset, i's indirect utility is as follows:

$$V_i(a, p, m_i) = \mathbf{E}\, U_i((1 + \rho)w_{i0} + a(x - m_i - (1 + \rho)p)), \tag{1}$$

where $w_{i0} = w_{ic} + pw_{ia}$ is the i's initial wealth. Agent i's demand in the risky asset is

$$D_i(p, m_i) = \arg\max_a V_i(a, p, m_i),$$

where the maximum is taken over all $a \geq 0$ that satisfy the budget constraint $pa \leq w_{i0}$. Since V_i is concave in a, and $D_i(p, m_i)$ is well defined and unique for every m_i and p.

In the following proposition, we use a technical assumption which imposes that for all terminal payoffs w which can be achieved with positive probability, $R_i^r(w) \leq 1$. To ensure this, we introduce an upper bound on terminal wealth[4]

[4] The first component of the maximum defining K is an upper bound on terminal wealth when the budget is entirely used to buy the risky asset. The expressions $(x_{\max} - m_i)$ is the maximum terminal wealth per unit of risky asset. This is then multiplied by an upper bound of units which might be bought. The second component relates to an upper bound on terminal wealth when only the riskless asset is consumed. The expectation in this expression is an upper bound on the value of the risky asset.

$$K = \text{Max}\left[(x_{\max} - m_i)\left(\frac{w_{ic}(1+\rho)}{x_{\min} - m_i} + w_{ia}\right), (1+\rho)w_{ic} + w_{ia}\mathbf{E}[x - m_i]\right]..$$

Useful properties of D_i are summarized below.

Proposition 1

1. $D_i(p, m_i) = \frac{w_{i0}}{p}$ *if* $p \leq \frac{x_{\min} - m_i}{1+\rho}$
2. $D_i(p, m_i) = 0$ *if* $p \geq \frac{\mathbf{E}x - m_i}{1+\rho}$ *and* $D_i(p, m_i) > 0$ *otherwise. Assume for all* $w \leq K$, $R_i^r(w) \leq 1$,
3. *the left-hand and right-hand derivatives of* D_i *exist at all* $p > 0$,
4. $D_i(p, m_i)$ *is decreasing in* p *for* $p < \frac{\mathbf{E}x - m_i}{1+\rho}$.

The proof of Proposition 1 can be found in Appendix 2.

Relative risk aversion is bounded by 1 in several cases of interest, including logarithmic utility functions. It is also consistent with Chetty (2006)'s estimate of 0.71 for the mean relative risk aversion in the population.

Our upper bound K is compatible with a CARA assumption we use in the analysis of Sect. 3.1. The more general DARA case analyzed in Sect. 3.2 is compatible with relative risk aversion bounded by 1 at all wealth levels. In the rest of the paper, we assume that the conclusions of Proposition 1 hold, which subsumes, but is not limited to the case of relative risk aversion bounded by 1 on a large enough range of wealth.

2.2 Market equilibrium

The total demand when management costs for all agents are represented by a vector $m = (m_i)_i$ is given by $D(p, m) = \sum_i D_i(p, m_i)$. The total supply of risky asset in the economy is $W_a = \sum_i w_{i,a}$ which we assume to be positive. An equilibrium price is a price p at which total demand equals supply, it thus satisfies

$$D(p, m) = W_a.. \tag{2}$$

Theorem 1 *There exists a unique market equilibrium price* $p(m)$ *for every vector* $m = (m_i)_i$ *of management costs.*

Proof of Theorem 1 We have established that D is continuous, equal to 0 for p large enough and going to ∞ for $p \to 0$. Locally, it is either equal to 0 (when $p \geq \frac{\mathbf{E}(x - m_i)}{1+\rho}$ for all i) or decreasing in p. By the intermediate value Theorem, there is a unique value of p such that $D(p, m) = W_a$.

3 Impact of management costs on market prices

Here, we focus on a comparison between equilibrium prices with management costs of zero and non-zero management costs via a first-order approximation[5] via the partial derivatives of $p(m)$ at $m = \mathbf{0}$, where $\mathbf{0}$ is the vector of zeros $(0, \dots, 0) \in \mathbb{R}^I$:

$$p(m) - p(\mathbf{0}) \sim \sum_i m_i \frac{\partial p(m)}{\partial m_i}|_{m=\mathbf{0}}. \tag{3}$$

To understand how management costs impact equilibrium prices p, we first study their impact on demand. Note that management costs impact demand both directly, since they impact net cash flows from the risky asse, and indirectly, through wealth effects. In fact, since management costs impact prices, they also impact agents' initial wealth $w_{i0} = w_{ic} + pw_{ia}$, hence their attitude toward risk.

To make the wealth effects explicit, we let $D_i(w_{ic}, w_{ia}, p, m_i)$ be the demand for the risky asset of agent i with initial allocation w_{ic} in cash and w_{ia} in the risky asset. $D_i(w_{ic}, w_{ia}, p, m_i)$ depends on w_{ic} and w_{ia} only through initial wealth $w_{i0} = w_{ic} + pw_{ia}$. We assume that we are in the economically relevant case in which the solution of the portolio optimization problem is interior: $D_i(w_{ic}, w_{ia}, p, m_i)$ is neither 0, nor the maximal amount $\frac{w_{i0}}{p}$. In this case, the solution of the portfolio optimization problem is then the same as the classical one in which there is no constraint on the amount of risky asset invested, there is no budget constraint and the investor could short the risky asset.

Classical results in choice under uncertainty (Arrow 1965; Pratt 1964) then show that under constant absolute risk aversion (CARA), $D_i(w_{ic}, w_{ia}, p, m_i)$ is then constant, whereas with decreasing absolute risk aversion (DARA) preferences, it is nondecreasing in initial wealth (see also proposition 8 in Gollier 2001).

The following proposition relates demand with and without management costs.

$$D_i(w_{ic}, w_{ia}, p, m_i) = D_i(w_{ic} - \frac{m_i}{1+\rho} w_{ia}, w_{ia}, p + \frac{m_i}{1+\rho}, 0).. \tag{4}$$

Proposition 2

Equation (4) has a natural interpretation. Management costs make the risky asset more expensive by $\frac{m_i}{1+\rho}$, this is the direct effect. But this apparent price increase is only faced by i when acquiring units of the risky asset. It does not affect the value of her initial endowment. Hence, to keep initial wealth constant, we need to deflate initial wealth by $\frac{m_i}{1+\rho} w_{ia}$.

Proof of Proposition 2 We compare demand at price p and at price $p' = p + \frac{m_i}{1+\rho}$. We express i's final wealth as

$$(1+\rho)w_{i0} + a(x - m_i - (1+\rho)p) = (1+\rho)w_{i0} + a(x - (1+\rho)p') \quad .$$

[5] Investment management costs reported by institutional investors at a portfolio level are typically relatively small, of the order of 3 to 30bps depending on the mix of asset classes.

Note that $w_{i0} = w_{ic} + pw_{ia} = w'_{ic} + p'w_{ia}$, with $w'_{ic} = w_{ic} - \frac{m_i}{1+\rho} w_{ia}$. This means that initial wealth holding w_{ic} in cash and w_{ia} in risky asset at price p is the same as with w'_{ic} in cash and w_{ia} in risky aset at price p'. It follows that the expected utility $V_i(w_{ic}, w_{ia}, a, p, m_i)$ from holding a units of the asset satisfies

$$
\begin{aligned}
V_i(w_{ic}, w_{ia}, a, p, m_i) &= \mathbf{E}U_i((1+\rho)w_{i0} + a(x - m_i - (1+\rho)p))) \\
&= \mathbf{E}U_i((1+\rho)w_{i0} + a(x - (1+\rho)p'))) \\
&= V_i(w_{ic} - \frac{m_i}{1+\rho} w_{ia}, w_{ia}, a, p', 0) \\
&= V_i(w_{ic} - \frac{m_i}{1+\rho} w_{ia}, w_{ia}, a, p + \frac{m_i}{1+\rho}, 0)
\end{aligned}
.
$$

Hence, the demand that maximizes the first expression in the series of equalities also maximizes the last one.

For simplification of exposition, we first analyze the case in which agents exhibit CARA preferences, in which case wealth effects are absent. We then show that, in the more general case where agents exhibit DARA preferences, the impact of management cost on demand and on prices can only be larger than in the CARA case.

3.1 Analysis absent wealth effects

In the reminder of this section, we assume that the agent invests neither 0 nor their whole wealth in the risky asset, in which case, as noted before, demand for the risky asset is independent of initial wealth and will thus be denoted $D_i(p, m_i)$. The more general case is treated in Sect. 3.2.

We estimate the impact of (small) management costs on the equilibrium price $p(m)$, around $m_i = 0$. In this case, Eq. (4) becomes

$$D_i(p, m_i) = D_i(p + \frac{m_i}{1+\rho}, 0). \tag{5}$$

The elasticity of demand for agent i is given by

$$e_i(p;m) = -\frac{\partial \log D_i}{\partial \log p} = -\frac{p}{D_i} \frac{\partial D_i}{\partial p}$$

where all derivatives are taken to be right-hand derivatives.

Differentiating (5) with respect to m_i at $m = \mathbf{0}$ gives

$$\frac{\partial D_i}{\partial m_i}|_{m_i=0} = \frac{1}{1+\rho} \frac{\partial D_i}{\partial p}|_{m_i=0} = -\frac{D_i}{p(1+\rho)} e_i, \quad , \tag{6}$$

where e_i is shorthand for $e_i(p;\mathbf{0})$, $\frac{\partial D_i}{\partial m_i}$ is a right-hand derivative, and $\frac{\partial D_i}{\partial p}$ a left-hand one.

By differentiating (2) and applying the implicit function theorem we obtain

$$\frac{\partial p(m)}{\partial m_i} = -\frac{\frac{\partial D_i}{\partial m_i}}{\sum_j \frac{\partial D_j}{\partial p}}, \tag{7}$$

where $\frac{\partial p(m)}{\partial m_i}$ and $\frac{\partial D_i}{\partial m_i}$ are taken as the right-hand derivatives, and $\frac{\partial D_j}{\partial p}$ as the left-hand derivative. The interpretation of the negative sign of the partial derivative is that since an increase in m_i would lead to a decrease in D for a constant p, it is compensated for by a decrease in p.

Finally, we combine (6) and (7) and obtain

$$\frac{\partial p(m)}{\partial m_i}|_{m=\mathbf{0}} = -\frac{1}{1+\rho}\frac{D_i e_i}{\sum_j D_j e_j}.. \tag{8}$$

By reporting (8) into (3), we obtain the fundamental expression:

$$p(\mathbf{0}) \sim p(m) + \frac{1}{1+\rho}\frac{\sum_i D_i e_i m_i}{\sum_i D_i e_i}.. \tag{9}$$

We let the *cost correction term* to be

$$\bar{m} = \frac{\sum_i D_i e_i m_i}{\sum_i D_i e_i}.. \tag{10}$$

The price $p(m)$ is the observed market price. The cost correction term is the average of agents' costs weighted by their portfolio sizes D_i and elasticities e_i. Demands D_i and associated costs m_i can to some extent be inferred from financial reporting and surveys. Elasticities e_i are not directly observed, but may either be econometrically estimated or inferred through agents' risk attitudes.

If we make the simplifying assumption that all agents have the same price elasticity in the risky asset we obtain that

$$\bar{m} = \frac{\sum_i D_i m_i}{\sum_i D_i}, \quad . \tag{11}$$

Hence, the cost correction term is the average of costs weighted by portfolio sizes, which can be estimated from reported data.

Differences in management costs, at least across institutional investors, should be an order of magnitude smaller than the average management costs. In case it is considered reasonable to assume all agents have the same management costs, we have

$$\bar{m} = m_i, \tag{12}$$

for every agent i.

3.2 Wealth effects

We now analyze the more general case in which U_i exhibits (weakly) decreasing absolute risk aversion (DARA) and where the budget constraint may be binding or not. We write $D_i(p, m_i) = D_i(w_{i0}, p, m_i)$ where $w_{i0} = w_{ic} + pw_{ia}$ is i's initial wealth. With DARA utility, demand for the risky asset is non-decreasing in the agent's initial wealth. In fact, we must be in one of three cases (1) the constraint $D_i(p, m_i) \geq 0$ is binding, in which case it is locally constant and (2) the budget constraint is binding and then $D_i(p, m_i) = \frac{w_{i0}}{p}$ is locally increasing in initial wealth w_{i0}, or none of them is binding and $D_i(p, m_i)$ non-increasing following classical results. We thus have

$$\frac{\partial D_i}{\partial w_{i0}}(w_{i0}, p, m_i) \geq 0.. \tag{13}$$

Differentiating (4) with respect to m_i at $m_i = 0$ gives

$$\begin{aligned}\frac{\partial D_i}{\partial m_i}|_{m_i=0} &= -\frac{w_{ia}}{1+\rho}\frac{\partial D_i}{\partial w_{i0}}|_{m_i=0} - \frac{D_i e_i}{p(1+\rho)} \\ &\leq -\frac{D_i e_i}{p(1+\rho)}|_{m_i=0}\end{aligned}.$$

Note that (7) is still valid as it is derived from (2) which expresses equality of supply and demand and does not hinge on agents' preferences or interior solutions. We thus combine the last inequality with (7) and obtain

$$\frac{\partial p(m)}{\partial m_i}|_{m=\mathbf{0}} \leq -\frac{1}{1+\rho}\frac{D_i e_i m_i}{\sum_j D_i e_i}.$$

Finally, we report this last result into (3), which gives us

$$p(\mathbf{0}) \geq p(m) + \frac{1}{1+\rho}\frac{\sum_i D_i e_i m_i}{\sum_i D_i e_i},, \tag{14}$$

where the inequality hold asymptotically for small m. In practice, this means that the cost correction term obtained without wealth effects is a lower bound of its true value, if wealth effects are taken into account.

This has a natural economic interpretation. In fact, by the direct effect, an increase in management costs make the risky asset less attractive, which in turn reduces its price. Moreover, by the wealth effect, when the price of the risky asset goes down, the agent's initial wealth goes down as well, and so its aversion to risk increases, leading to lower demand for the risky asset, lowering its price again. As we see, wealth effects amplify the direct effect, hence lead to a larger cost correction term.

4 Implications for insurance accounting

In this section, our examination begins with an analysis of the valuation approach currently implemented under Solvency II to incorporate investment management costs. We demonstrate that this methodology results in cost double counting and propose a suitable correction. These findings are then utilized to gauge the scale of cost double counting in insurance balance sheets. We discuss the specific balance sheet items affected by this double counting and estimate its impact on solvency ratios. Lastly, we delve into the distorted incentives resulting from cost double counting and explore their implications for investment strategies.

4.1 Double counting of management costs under Solvency II

Under Solvency II, SST and BSCR regulations, insurers are required to value assets and liabilities at fair value.[6] Traditional life insurance liabilities depend on the cash flows of the investments which cover these liabilities. To value those liabilities, the Solvency II delegate regulation, EU Parliament (2015), requires insurers to use an arbitrage-free and market-consistent model,[7] and moreover to consider for their valuation their own management expenses.[8] Similar requirements exist for SST in FINMA (2020) and BSCR in Bermuda Monetary Authority (2011).[9]

In practice, insurers implement those requirements via a (potentially stochastic) projection of asset and liability cash flows and by discounting them with a risk-free interest rate curve. The discounting process is constructed so that the gross discounted investment cash flows correspond to the market value of the investments. Hence, those assets are valued while implicitly assuming a market where investors bear no management costs. In this case, its market price would be the equilibrium price with management costs of zero: $p(\mathbf{0})$. In parallel, the present value of the future investment management costs are added to the liability side.

Insurers value their liabilities, which typically depend on their investments and investment strategy with a model which values the cash flow of one unit of the risky asset they hold, including their management costs, as

$$v_i = p(\mathbf{0}) - \frac{m_i}{1+\rho}. \quad . \tag{15}$$

[6] See EU Parliament (2009), (45), (54), §75, 76.3 and EU Parliament (2015) §10 for Solvency II.

[7] See ,e.g., §22.3: "Where insurance and reinsurance undertakings use a model to produce projections of future financial market parameters, it shall comply with all of the following requirements: (a) it generates asset prices that are consistent with asset prices observed in financial markets; (b) it assumes no arbitrage opportunity."

[8] See §31: "A cash flow projection used to calculate best estimates [⋯] takes into account various expenses, including investment management expenses."

[9] See pages 3, 13 and 20 for FINMA (2020) and page 152 paragraph 8.(b)(ii) and paragraph 15 for Bermuda Monetary Authority (2011). We focus here on capital metrics according to the strictest metrics, in terms of Solvency II the Solvency II balance sheet without transitional measures, volatility and matching adjustment, or in terms of BSCR assuming the equivalent so-called scenario based approach.

When a stochastic projection is required, such a valuation is usually obtained as the expectation under risk-neutral probabilities. Construction of risk-neutral probabilities in the presence of management costs is discussed in Appendix 1.

It should be acknowledged that investment management costs are not confined to the insurance company alone (which might constitute a small segment of the wider market). Other insurance companies and, more broadly, all other investors also incur these costs. As such, $p(\mathbf{0})$, the equilibrium price devoid of management costs, is not observed in reality. What is observed is $p(m)$, the equilibrium price that incorporates management costs of m. This distinction between the observed prices $p(m)$ and the management cost-free prices $p(\mathbf{0})$ is overlooked by insurers. They inaccurately use $p(m)$ in their models as a substitute for $p(\mathbf{0})$. This misunderstanding leads to a flawed valuation of one unit of the risky asset in practice, as illustrated per the formula:

$$\vartheta_i = p(m) - \frac{m_i}{1+\rho}. \quad . \tag{16}$$

To see that this leads to some double counting of costs it helps to analyze the special case of CARA preferences whereby (9), the above equation can be rewritten as follows:

$$\vartheta_i = p(m) - \frac{m_i}{1+\rho} \sim p(\mathbf{0}) - \frac{1}{1+\rho}\left(\bar{m} + m_i\right). \quad . \tag{17}$$

This final expression demonstrates the occurrence of a double counting of management costs. Firstly, these costs are implicitly factored into the market price, represented by their weighted average $\frac{\bar{m}}{1+\rho}$. Subsequently, individual costs are deducted once more through $\frac{m_i}{1+\rho}$. This double counting of management costs results in an exaggeration of management costs in the present valuation approach.

By combining (9) and (15), we obtain that the correct valuation with CARA preferences is given by

$$v_i \sim p(m) + \frac{1}{1+\rho}\left(\bar{m} - m_i\right). \quad . \tag{18}$$

With wealth effects, by (14) we obtain that the correct valuation is as follows:

$$v_i \geq p(m) + \frac{1}{1+\rho}\left(\bar{m} - m_i\right), \tag{19}$$

and therefore, the correction term is even larger than in the CARA case.

For the remainder of this section we focus on CARA preferences. Current Solvency II valuations per unit of risky asset have to be corrected by the amount:

$$v_i - \vartheta_i \sim \frac{1}{1+\rho}\bar{m}. \quad . \tag{20}$$

In practical terms, the disparity in management costs borne by reasonably efficient investors is substantially smaller than the average management costs. Such minor differences may be attributable to variances in efficiency, strategy, or the quality of asset management. Consequently, the term $\bar{m} - m_i$ should be relatively small when compared to $\bar{m}$. This suggests that deviations in individual management costs from the market average are likely to be minimal and should not significantly alter the overall cost calculation.

Therefore, without any explicit evidence of markedly superior or inferior efficiency for specific insurers compared to the market average, it may be reasonable to postulate that costs, denoted by m_i, are identical for all market participants. This would mean that $m_i = \bar{m}$ for every individual i, regardless of their respective elasticities and portfolio sizes. The valuation represented by Eq. (18) would then simplify to

$$v_i \sim p(m). \quad . \tag{21}$$

In such a scenario, investment management costs would effectively be entirely factored in market prices. Consequently, the provision of $\frac{m_i}{1+\rho}$ in the Solvency II balance sheet would be entirely superfluous.

4.2 Quantitative impact

The analysis provided herein serves two primary roles within the context of Solvency II. The first involves estimating and analyzing the impact of current double counting at the aggregate insurance industry level and drawing policy recommendations from these findings. The second concerns individual corrections of double counting at the company level. It is important to note that a precise estimation of double counting at the firm level may require considering the investment portfolio and strategy toward risky assets. Following this, it would be necessary to estimate the broader market's costs for managing such a portfolio. With some simplifying assumptions, one might then calculate the cost correction term $\bar{m}$ per unit of risky asset, as defined in (10), and ascertain the deviation of the company's own costs m_i as the amount of costs to be provisioned for.

In this section, we focus on the more policy relevant question of estimating the magnitude of double counting at the aggregate level. The objective of this study is not to attain a high degree of precision, but rather to ascertain whether double counting might have a significant enough impact on balance sheets to warrant attention from policy-makers. To achieve this, we focus on two key aspects: First, the volume of cost double counting present in European insurance balance sheets and second, the specific impact of a correction of double accounting on different items of the balance sheet.

4.2.1 Estimation of the magnitude of double counting

To estimate the magnitude of double counting, we apply our model with a single risky asset, to insurers who typically invest in multiple risky asset classes. The mix changes from one insurer to the other and the focus on different risky asset classes varies also from one European country to another.

We will therefore make some simplifying assumptions. We will ignore wealth effects, keeping in mind that double counting would be exacerbated by wealth effects. We assume that insurers within a country invest into a single risky asset. One may also think of a fixed risky asset mix close enough across insurers. For our double counting estimate, one may think of applying the model country by country or assuming that across countries only one risky asset is considered.

We will also assume that all investors have either the same or close enough price elasticities. This allows us to apply equation (11). Alternatively one could assume that management costs within each country we consider are identical. This would allow for applying equation (12). If neither is satisfied, this would imply some bias in our estimations, but this should be seen in the context of us being primarily interested in an order of magnitude.

We will focus only on the insurance sector which may imply some other bias. Insurance companies are substantial institutional investors. They should be able to obtain the services of portfolio managers at costs well below that of retail investors and approximately in line with, or even lower than, other institutional investors. This includes entities such as family offices and corporate pension funds. Consequently, the average market management cost $\bar{m}$ should be at least equal to the cost borne by individual insurance companies on average. This implies that the amount of double counting for insurance companies, as a whole, should be at the least the sum of the management cost provisioned for in the Solvency II balance sheets. The bias induced by focusing on insurance company data only should rather underestimate the full amount of double counting.

We do not have any direct access to cost provisions under Solvency II, but we can estimate those at the aggregate level based on public data. Let us hence denote

- Inv the total value of investments managed by the insurance industry.
- Du the modified duration of the insurance liabilities, i.e., the discounted average time over which the assets need to be held to back liabilities;
- μ the average annual investment management costs of insurers (or according to the context of insurers in a given country) as a percentage of their investments and across their entire portfolio including the risky and riskless asset,

Assuming costs as percentage of the investments are stable throughout the life of the liabilties, the Euro amount of management costs generated can be estimated through the formula:

$$M = Inv \times Du \times \mu \quad . \tag{22}$$

Linking those parameters to our model, we should have the following interpretations

- $Inv = \sum_{i \in I} \left(w_{ic} + p w_{ia} \right)$ where we have only retained insurers as agents I
- $M = \frac{1}{1+\rho} \sum_{i \in I} m_i D_i$.

The present value of the average management costs per unit of risky asset is

$$\frac{M}{\sum_{i\in I} D_i} = \frac{1}{1+\rho} \frac{\sum_{i\in I} m_i D_i}{\sum_{i\in I} D_i}. \tag{23}$$

If either elasticities are constant or if all m_i are equal, then we can apply equations (11) or (12), leading to

$$\frac{M}{\sum_{i\in I} D_i} = \frac{\bar{m}}{1+\rho}.. \tag{24}$$

This corresponds to the amount of double counting per unit invested. The total amount of double counting is then given by

$$\sum_{i\in I} D_i \frac{\bar{m}}{1+\rho} = M.. \tag{25}$$

We base our estimation of the liability duration on EIOPA (2019b).[10] The market value of investments managed by European insurers are reported in EIOPA's insurance statistics database,[11] which also contains the reported statutory investment management costs *SC*. These will form the foundation to estimate μ. While there might be differences between costs from statutory reporting and costs provisioned under Solvency II, the statutory costs are a natural starting point and the only publicly available data that we are aware of. Generally, costs that are incurred off-balance sheet are not included in *SC*. For instance, the costs of managing the investments of a fund are typically borne by the funds themselves and are not part of the reported statutory costs. In general, we assume

$$\mu = \frac{SC}{Inv \times (1 - CIU)}, \tag{26}$$

where *Inv* corresponds to the market value of investments and *CIU* corresponds to the share of investments held via collective investment undertakings (funds).

Table 1 presents data on reported investments, investment management costs, and the liability duration by country. From this, we deduce the estimation of double counting.

The resulting figures are designed to provide an estimate of the order of magnitude of double counting under the strictest Solvency II measure, without transitional measures, volatility, and matching adjustments.[12] As Table 1 indicates, the total impact on the EEA is estimated at an order of magnitude of € 130 bn, with the most affected countries in absolute terms being Germany (€ 67 bn) and France (€ 30 bn).

It is noteworthy that the reported costs vary considerably between countries. We do not possess sufficiently detailed data to scrutinize these differences closely. The costs difference should be partly explained by different reporting standards across

[10] See page 38.

[11] Available on https://www.eiopa.europa.eu.

[12] The Solvency II balance sheet with matching adjustment (which is used mainly in the UK and to some extent in Spain) contains less double counting of costs than the strictest Solvency II metric on which we focus.

Table 1 2021 investments and costs

Country	*Inv* (€ bn)	*SC* (€ bn)	*CIU*	μ (bps)	*Du*	$Du \times \mu$ (%)	*M* (€ bn)
DE	2265	2.4	32%	16	19.4	3.0	67
DK	306	0.6	–	20	14.1	2.9	9
FR	2371	2.0	20%	11	11.8	1.3	30
IT	817	0.9	13%	12	9	1.1	9
NL	404	0.4	7%	11	13.4	1.5	6
EEA	7547	9.0	21%	15	11.9	1.8	134

Inv = investments - derivatives + loans and mortgages, SC = statutory costs, CIU = share of collective investment undertakings, $\mu = \frac{SC}{I\times(1-CIU)}$ = costs adjusted for costs not reported for CIU, except for Denmark where we set $\mu = \frac{SC}{Inv}$. Du = liability duration, $Du \times \mu$ = correction term in percentage points of investments, $M = Inv \cdot Du \cdot \mu$ = absolute correction term versus a naive modeling approach

Europe. For instance, Danish life and pension insurers report consolidated investment management costs. Since the investments that are more expensive to manage are typically not held directly on the balance sheet, this should account for at least part of the higher reported costs for Denmark. There is also a varying focus on more or less costly to manage asset classes country by country.[13]

The tendency for expensive-to-manage assets to be held indirectly, with their costs not appearing in statutory accounting reports, may lead to some systematic underestimation of costs. This is based on our assumption that costs charged to funds align with statutory costs, which mainly cover investments held directly on the balance sheet.

Conversely, many insurers do not account for the maintenance costs of real estate investments in their Solvency II provisions. Consequently, our model may overestimate cost double counting for this asset class, particularly when such investments are held directly on the balance sheet and form part of the statutory cost reporting.[14]

4.2.2 Impact on balance sheets

Here, we examine the impact of eliminating double counting on Solvency II balance sheet items as well as on solvency ratios.

[13] Based on discussions with many stakeholders across countries, asset management costs by type of mandate/asset class across Europe seem fairly similar and could not at all explain the differences in reported investment management costs. Moreover, while we lack publicly available data, our anecdotal evidence hints to material differences across countries regarding the typical focus on costly to manage asset classes.

[14] While it is common practice to apply cost double counting broadly across all asset classes except real estate, this does not hold true for every individual insurance company across all countries and asset classes. In a consultation paper, EIOPA (2021) (see pages 12 and 13) suggests that all expenses should be taken into account in line with the strategy, at least for investments backing technical provisions and investments backing the solvency requirement. Hence, only management costs for investments backing excess capital could be ignored, if those considerations are implemented.

Eliminating the double counting of costs results in an additional cash flow in the Solvency II models. Under Solvency II, future cash flows are valued post-tax. For non-life insurers, the additional cash flow is to be divided between future tax payments, represented by the deferred tax liability in the Solvency II balance sheet, and payments to shareholders, which correspond to Solvency II own funds.

In a stress scenario that determines solvency capital requirements (SCR), typically less income is generated, and the amount of taxes paid decreases. These reduced tax payments absorb some of the costs of the shock and correspond to the loss-absorbing capacity of deferred taxes (LAC DT) in the Solvency II balance sheet. Consequently, for non-life insurers, eliminating cost double counting should influence both the numerator and the denominator of the capital ratio, increasing own funds through payments to shareholders and decreasing SCR through deferred tax payments which may absorb losses.

Life insurers manage most of European insurance assets. In the Solvency II balance sheet projection of traditional with-profit life insurers, any additional cash flow would initially flow through profit sharing. In many countries, around 85% of income generated must be allocated to policyholders, as is the case in France and Italy, while in Germany, the figure is 90%. As such, approximately 85% of the modeled extra cash flows should flow into future discretionary benefits (FDB). In stress scenarios like those determining the SCR, the base case projected FDB can be reduced to absorb a shock. Thus, the FDB, which should be increased via a correction for the double counting of costs, could to some extent be used as the loss-absorbing capacity of technical provisions (LAC TP), thereby reducing the SCR. The share of FDB used as LAC TP can vary greatly across insurers and largely depends on modeling choices, which should reflect the company's strategy. We assume that at least 70% of the additional FDB could be used as LAC TP, thereby reducing the SCR by that amount. This assumption aligns with the Solvency II reporting of German and French life insurers over the past years.

Assuming for both life and non-life insurers that on average (i) about 80% of the eliminated double counting flows to FDB, (ii) 70% of FDB is used as LAC TP, (iii) the tax rate is 30%, and (iv) 70% of deferred taxes are used as LAC DT, we can now make a rough estimation of the impact of eliminating cost double counting in the European Economic Area (EEA):

- Increase of own funds by 20% × 70% × € 134 billion = € 19 billion,
- Increase of deferred tax liabilities by 20% × 30% × € 134 billion = € 8 billion,
- Increase of FDB by 80% × € 134 billion = € 107 billion,
- Decrease of SCR by 70% × (€ 8 billion + € 107 billion) = € 81 billion.

According to EIOPA's insurance statistics under the Solvency II metric, without transitionals, volatility, and matching adjustments, the combined eligible own funds and solvency capital requirements for the EEA insurance industry for 2021 were, respectively, € 973 billion and € 459 billion, leading to an average Solvency ratio of 212%.

Increasing own funds by € 19 bn and decreasing SCR by € 81 bn would increase the Solvency ratio by 50 percentage points to 262%. However, due to limitations in publicly available data, our estimate of the potential impact on the average Solvency ratio should be viewed as a rough approximation. Moreover, some insurers in some countries do not systematically apply cost double counting across all asset classes.

4.3 Distortion of incentives

In this section, we discuss the consequences of double counting on incentives and asset allocation.

The double counting of costs leads to an unjustified capital buffer as a result of an improper application of Solvency II principles. This extra accounting burden, by raising capital requirements, may diminish the appeal of capital to the insurance industry. However, prudential requirements are the outcome of a process that takes into account the aggregate requirements. Any gains from eliminating double counting might be interpreted as a loosening of prudential rules which the regulator may subsequently offset by a tightening of the requirements on other aspects.

In our view, the more critical consequence of double counting of management costs is the likely impact on capital allocation through distorted incentives. To understand this, it is useful to illustrate how investment management costs can vary significantly with the investment strategy. Table 2 is based on European pension funds with more than € 2,000 billion of assets under management, as reported by Beath and Flynn (2018).[15]

Given the duration of life insurance liabilities, expensive-to-manage assets used to back these liabilities lead to a massive burden on the Solvency II balance sheet when the costs are imputed twice.[16]

Therefore, cost double counting pushes life insurance companies to shy away from assets with high management costs which are precisely the types of assets that policyholders often cannot access efficiently themselves. This challenges the role life insurers should play in organizing long-term savings and can be detrimental to policyholders, insurance company shareholders, and the role of insurance companies in financing the economy.

5 Conclusion and policy implications

In a two-period general equilibrium model, we estimated the impact of management costs on asset prices. These assets could be listed securities, private assets like loans, private equity, actively or passively managed portfolios, funds, and so on.

[15] For a set of US pension funds with $ 2, 900 bn of assets under management, Beath and Flynn (2020) find costs for US fixed income ranging from 9 bps (other) to 18 bps (long-duration bonds). For Europe, we are not aware of any publicly available survey with granular data for management costs by fixed income subclass.

[16] Insurers counting costs twice may have incentives to assume in their Solvency II projections that they switch from a costly asset allocation to very cheap to manage investments, well before liabilities roll off. This would reduce the present value of overall modeled costs, at the cost of a more complex model. Such an approach also changes rightly or wrongly the valuation of the policyholder's options and guarantees. Eliminating double counting would eliminate these incentives.

Table 2 CEM benchmarking 2018

Asset class	Dutch	Other EU	UK
Public equity	7	12	11
Private equity	454	382	415
Fixed income	6	4	5
Hedge funds	261	258	227
Listed real estate	28	24	78
Unlisted real estate	114	46	69
Infrastructure	159	150	187
Other	31	64	100

Costs are expressed in basis points (bps)

We show that deducting management costs from the valuation of a portfolio, based on observed or estimated prices, results in a double counting of these management costs.

We suggest applying a correction term to model investment management costs in the Solvency II, SST, or BSCR balance sheets, specifically focusing on the strictest available Solvency II metric excluding transitionals, volatility, and matching adjustment. An analysis with any of these measures are beyond the scope of this document.

It should be noted that the implementation of IFRS 17 should also lead to cost double counting, albeit to a lesser extent.[17]

Correcting for double counting of management costs could have substantial implications in terms of available capital, capital management, and investment strategies. Possibly the most significant consequence of the current status quo is its impact on investment strategies by distorting investments away from investment alternatives that are costly to manage toward cheap ones.

We emphasize that a key value of pooled investment activities, like those offered by life insurance, is providing retail investors with access to investments that they ordinarily could not fully or even partially access, such as private debt, private equity, and infrastructure equity, among others. Status quo modeling penalizes such investments for the wrong reasons, i.e., the fact that they are more expensive to manage. This pushes insurers toward cheaper-to-manage asset classes, which also undermines the role the insurance sector can play in financing the economy as an investor well positioned to support relatively complex to manage long-term investments.

Implementing a correction could also mitigate the capital inefficiency of traditional with-profit business, which some insurers have put into run-off due to high capital consumption. In some cases, this might even enable some companies to redeploy capital.

[17] This is partly due to the fact that we expect the largest effect of an elimination of cost double counting to be on the SCR, which is not an IFRS concept. See in IASB (2021) explanations on *Cash flows within the contract boundary (paragraph 34)* the point B.65.(ka). The standard requires in §33 market-consistent modeling which may be achieved via “risk-neutral measurement techniques” according to B.77.

Finally, eliminating double counting could have an impact on mergers and acquisitions for traditional life insurance portfolios. Consolidators who acquire such portfolios would often shift the asset allocation to more alternative debt or more generally, to assets that are more expensive to manage. The adverse impact of the status quo modeling on the capital needed to support the book of business would then be exacerbated.

The current data does not allow for an exact determination of the market's weighted average management costs. However, it is reasonable to expect that differences in costs are much smaller in magnitude than absolute costs. A significant number of institutional investors disclose their investment management costs and consulting firms provide benchmarking services. These sources offer a more detailed understanding of a substantial range of investors' management costs. While this information can guide in identifying clear inefficiencies, minor variations in costs might be attributable to differences in the quality of investment strategies for specific market segments.

A strategy that incorporates differences in management costs, which are not directly related to either superior or inferior efficiency, could inadvertently disadvantage those operations that prioritize thoroughness and consequently incur higher costs. Conversely, it may unduly favor those that lack such diligence. This risk, however, must be balanced against the beneficial impact of obliging firms to assess the present value of costs arising from inefficiencies that haven't been addressed.

A sensible approach to a correction of the double accounting might involve an initial assessment to identify any significant cost discrepancies that can be traced back to inefficiencies. If such discrepancies are detected, it may be appropriate to calculate their magnitude and incorporate the corresponding present value into the liability segment of the insurance balance sheet. On the other hand, if the costs associated with an investment strategy appear to align with market averages, it could be most pragmatic to assume uniformity in investors' costs. As indicated by formula (21), this assumption would result in the negation of investment management costs in the valuation process.

Appendix 1: Risk-neutral probabilities

In practice, Solvency II balance sheets are often valued based on risk-neutral simulations.[18] For the practitioner, it may hence be useful to discuss how risk-neutral probabilities could be constructed while taking into account costs.

[18] According to EIOPA (2019a) §3.3.93, *"for valuing the best estimate for non-unconditional benefits, a stochastic simulation approach would consist of an appropriate market-consistent asset model for projections of risk-neutral returns"*

Following Drèze (1970), Ross (1977) and Cox and Ross (1976), risk-neutral probabilities[19] are probabilities Q over states of nature such that the price of every security X with payoff $X(s)$ in state s can be computed as follows:

$$p_X = \frac{1}{1+\rho} \int_s X(s) dQ(s).. \tag{27}$$

Risk-neutral probabilities exist under the absence of arbitrage opportunity. They are unique if furthermore, markets are complete. It should be noted that these probabilities *do not* represent objective probabilities of events, like probabilities over coin flips, but merely a convenient pricing instrument.

It follows from (27) that, under the risk-neutral probability asset, the return every asset is the same in every state and is equal to the risk-free return.

In our economy, each state s is associated with a payoff from the risky asset. There are two assets: the risk-free one which pays $1+\rho$ in every state of s, and the risky asset, which pays $X(s)$ in state s (where the state space S, P is the underlying probability space).

As markets are not necessarily complete, the risk-neutral probability measures are not necessarily unique. Any probability distribution Q that satisfies (27) for both the riskless and the risky assets is a risk-neutral probability.

Since the price $p(m)$ of the risky asset depends on management costs m, so do the corresponding risk-neutral probabilities. Therefore, an appropriate computation of risk-neutral probabilities should take into account cost considerations. If we denote by Q^m a risk-neutral probability when management costs are m, we have

$$p(m) = \frac{1}{1+\rho} \mathbf{E}_{Q^m} X, \tag{28}$$

and in particular

$$p(\mathbf{0}) = \frac{1}{1+\rho} \mathbf{E}_{Q^0} X.. \tag{29}$$

In practice, the relationships (27) are used to derive risk-neutral probabilities, under which classes of assets can be priced. Note that here we presented what we can call *gross cash flow risk-neutral probabilities*. They value the asset discounting the gross cash flows. Especially, when all costs are equal, it would be more convenient to work directly on net cash flows which would lead to a different pricing probability that one might call *net cash flow risk-neutral probabilities*. When there are no management costs, both concepts coincide.

To follow a valuation approach starting with an assumption of zero costs requires an evaluation of Q^0. In turn, the evaluation of Q^0 requires a prior estimation of $p(\mathbf{0})$, namely of the prices of assets absent management costs.

In order to properly calibrate a model of risk-neutral probabilities, one can therefore

[19] We consider here risk-neutral probabilities using the risk-free asset as numÃ©raire, but the discussion extends naturally to any other reference asset.

1. Estimate $p(\mathbf{0})$ from observed market prices and using formula (9),
2. Calibrate risk-neutral probabilities Q^0 from thus obtained prices,
3. Value portfolios under Q^0,
4. Subtract discounted management expenses from the obtained value, thus obtaining the net value v_i of the portfolio.

We then obtain

$$v_i = \frac{1}{1+\rho}\mathbf{E}_{Q(0)}X - \frac{m_i}{1+\rho} = p(\mathbf{0}) - \frac{m_i}{1+\rho} \tag{30}$$

which is the same as formula (15). It follows that v_i obtained using risk-neutral probabilities indeed coincides with formula (18). This concludes that our pricing method is in fact the same as under risk-neutral probabilities, once these probabilities are properly derived from the price system.

As an alternative to the method above, each firm can compute risk-neutral probabilities by equating asset prices with expected gross returns, thus using Q^m instead of Q^0 for the management of their assets. The advantage of this method is that Q^m is directly inferred from market prices, whereas Q^0 is not. If firms value portfolios based on Q^m, and subtract discounted management expenses from the obtained value, they need to also add back the discounted value of the cost correction term $\bar{m}$ to asset prices. In fact, the relationship

$$\mathbf{E}_{Q^m}X - \frac{m_i}{1+\rho} + \frac{\bar{m}}{1+\rho} = \mathbf{E}_{Q^0}X - \frac{m_i}{1+\rho} = p(\mathbf{0}) - \frac{m_i}{1+\rho}$$

shows that the pricing thus obtained coincides with (15).

Appendix B: Proof of Proposition 1

To prove the first point let $0 < p \leq \frac{x_{\min}-m_i}{1+\rho}$. Then $1+\rho \leq \frac{(x-m_i)}{p}$ a.s. and the inequality is strictly positive with positive probability. Hence for any $a < \frac{w_{i0}}{p}$ with probability 1 the payoff is lower or equal than allocating the entire budget, hence $\frac{w_{i0}}{p}$ to the risky asset, and with a positive probability it is strictly lower. Since U_i is strictly increasing we must have $\frac{w_{i0}}{p} = D_i(a,p)$.

The map V_i has well defined partial derivative V'_{ia} with respect to a given by

$$V'_{ia}(a,p) = \mathbf{E}\big[\big(x - m_i - (1+\rho)p\big)U'_i\big((1+\rho)w_{i0} + a(x - m_i - (1+\rho)p)\big)\big].$$

We will later use the fact that $V''_{iaa}(a,p) =$

$$\mathbf{E}\Big[\big(x - m_i - (1+\rho)p\big)^2 U''_i\big((1+\rho)w_{i0} + a(x - m_i - (1+\rho)p)\big)\Big] < 0.$$

Demand is given by

$$\begin{cases} D_i(p) = 0 & \text{if } V'_{ia}(0,p) \leq 0 \\ D_i(p) = \frac{w_{i0}}{p} & \text{if } V'_{ia}(\frac{w_{i0}}{p},p) \geq 0 \\ V'_{ia}(D_i(p)) = 0 & \text{otherwise .} \end{cases}.$$

For the second point, note that

$$\begin{aligned} V'_{ia}(0,p) =& \mathbf{E}\left[\left(x - m_i - (1+\rho)p\right) U'_i\left((1+\rho)w_{i0}\right)\right] \\ =& \mathbf{E}\left[x - m_i - (1+\rho)p\right] U'_i\left((1+\rho)w_{i0}\right), \end{aligned}$$

so that $V'_{ia}(0,p)$ has the same sign as $\mathbf{E}\left[x - m_i - (1+\rho)p\right]$, hence the result for the second point.

We now move to prove points 3 and 4. For $p < \frac{x_{\min}-m_i}{1+\rho}$ and $p > \frac{\mathbf{E}(x-m_i)}{1+\rho}$ we have the explicit expression of D_i and for those p the statements of point 3 and 4 hold obviously true.

It is hence enough to prove 3 and 4 for $p \in [\frac{x_{\min}-m_i}{1+\rho}, \frac{\mathbf{E}(x-m_i)}{1+\rho}]$ and for the below p is assumed to be within this interval. For a given price p, it will also be sufficient to consider risky asset quantities within the budget set hence $a \in [0, \frac{w_{i0}}{p}]$.

We first determine the sign of $V''_{iap}(a,p)$ in a similar way to Fishburn and Porter (1976):

$$\begin{aligned} V''_{iap}(a,p) =& -(1+\rho)\mathbf{E}\left[U'(w_{if}) - (x - m_i - (1+\rho)p)(w_{ia} - a)U''(w_{if})\right] \\ =& -(1+\rho)\mathbf{E}\left[\left(w_{if} + (x - m_i - (1+\rho)p)(w_{ia} - a)R^r_i(w_{if})\right)\frac{U'_i(w_{if})}{w_{if}}\right] \end{aligned}$$

with $w_{if} = (1+\rho)w_{i0} + a(x - m_i - (1+\rho)p)$. Given that $a \geq 0$, by the survival assumption, $w_{if} \geq (1+\rho)w_{i0} - a(1+\rho)p$ a.s. and then since $a \leq \frac{w_{i0}}{p}$ we have $w_{if} \geq 0$ a.s.

We will use that for all terminal payoffs w which can be achieved with positive probability for the considered prices, $R^r_i(w_{if}) \leq 1$. As we have assumed that $R^r_i(w) \leq 1$ for all $w \leq K$ with

$$K = \text{Max}\left[(x_{\max} - m_i)\left(\frac{w_{ic}(1+\rho)}{x_{\min} - m_i} + w_{ia}\right), (1+\rho)w_{ic} + w_{ia}\mathbf{E}[x - m_i]\right].$$

it is enough to prove that $w_{if} \leq K$.

For this, consider two cases: First assume $x - m_i \geq (1+\rho)p$, then w_{if} increases with a and therefore setting $a = \frac{w_{i0}}{p}$ we have

$$\begin{aligned} w_{if} \leq \frac{(x - m_i)w_{i0}}{p} &\leq (x_{\max} - m_i)\left(\frac{w_{ic}}{p} + w_{ia}\right) \\ &\leq (x_{\max} - m_i)\left(\frac{w_{ic}(1+\rho)}{x_{\min} - m_i} + w_{ia}\right) \end{aligned}.$$

Now assume $x - m_i \leq (1+\rho)p$ then w_{if} decreases with a and therefore setting $a = 0$ we have

$$w_{if} \leq (1+\rho)(w_{ic} + pw_{ia}) \leq (1+\rho)w_{ic} + \mathbf{E}(x - m_i)w_{ia}..$$

Hence, w_{if} is smaller than K which is the maximum of both terms.

Now, we show that $w_{if} + (x - m_i - (1+\rho)p)(w_{ia} - a)R_i^r(w_{if})$ is positive a.s.. For $R_i^r(w_{if}) = 0$ its value is $w_{if} \geq 0$ a.s. and for $R_i^r(w_{if}) = 1$ its value is $(1+\rho)w_{ic} + (x - m_i)w_{ia} > 0$ a.s. As $R_i^r(w_{if}) \in (0, 1]$ it lies between those two values. Hence $V''_{iap}(a, p) < 0$ for $p \in [\frac{x_{\min} - m_i}{1+\rho}, \frac{\mathbf{E}(x-m_i)}{1+\rho}]$ and $a \in [0, \frac{w_{i0}}{p}]$.

To complete the proof we will consider five cases for $p_0 \in [\frac{x_{\min} - m_i}{1+\rho}, \frac{\mathbf{E}(x-m_i)}{1+\rho}]$.

A. If $V'_{ia}(0, p_0) < 0$, then $D_i(p) = 0$ in a neighborhood of p_0 and hence by point 2, $p_0 > \frac{\mathbf{E}(x-m_i)}{1+\rho}$.
B. If $V'_{ia}(\frac{w_{i0}}{p_0}, p_0) > 0$ then $D_i(p) = \frac{w_{i0}}{p} = w_{ia} + \frac{w_{ic}}{p}$ in a neighborhood of p_0. Hence $D_i(p)$ is differentiable and strictly decreasing in a neighborhood of p_0, independently whether $p_0 < \frac{x_{\min} - m_i}{1+\rho}$ or not.
C. If $V'_{ia}(0, p_0) > 0$ and $V'_{ia}(\frac{w_{i0}}{p_0}, p_0) < 0$ then D_i is given by $V'_{ia}(D_i(p)) = 0$ in a neighborhood of p_0. By the implicit function theorem, D_i is differentiable at p_0 and decreasing as $D'_i(p_0) = -\frac{V''_{iap}}{V''_{iaa}}(D_i(p_0)) < 0$.
D. If $V'_{ia}(0, p_0) = 0$, then $D_i = 0$ and hence $p_0 \geq \frac{\mathbf{E}(x-m_i)}{1+\rho}$. We only need to analyze the case $p_0 = \frac{\mathbf{E}x-m_i}{1+\rho}$. Hence, $D_i = 0$ for $p \geq p_0$, so D_i has right-hand derivative 0 at p_0. Similarly D_i is given by $V'_{ia}(D_i(p)) = 0$ for $p < p_0$ close to p_0, so D_i has left-hand derivative $-\frac{V''_{iap}}{V''_{iaa}}(D_i(p_0)) < 0$ at p_0.
E. If $V'_{ia}(\frac{w_{i0}}{p_0}, p_0) = 0$, then by the implicit function theorem there exists a neighborhood of p_0 in which the solution $D_i^r(p)$ to $V'_{ia}(D_i(p), p) = 0$ exists, is unique, and is differentiable with derivative $-\frac{V''_{iap}}{V''_{iaa}}(D_i(p_0)) < 0$ at $p = p_0$. On this neighborhood, we have $D_i(p) = \min\{D_i^r(p), \frac{w_{i0}}{p}\}$. The derivative of $\frac{w_{i0}}{p}$ with respect to p exists and is negative. The left-hand and right-hand derivative of D_i at p_0 exists. The right-hand derivative is the minimum of the derivatives of $D_i^r(p)$ and $\frac{w_{i0}}{p}$ at $p = p_0$ and is hence negative. The left-hand derivative is the maximum of the derivatives of $D_i^r(p)$ and $\frac{w_{i0}}{p}$ at $p = p_0$ and hence also negative.

Acknowledgements We are grateful to Mario Müller for bringing the topic to our attention. This work evolved out of past advisory work for German insurance companies, and benefited from multiple discussions with insurance companies in Denmark, Germany, France, Italy, the Netherlands, Switzerland and the UK starting November 2019. Moreover, we are grateful for helpful discussions and comments from Jonathan Benchimol, Jürgen Bierbaum, Paul-Emmanuel Chouc, Helmut Gründl, Caroline Hillairet, Felix Kubler, Augustin Landier, Enrico Minelli, Herakles Polemarchakis, Romain Rancière, Oded Sarig, David Thesmar and detailed comments from two anonymous referees. All remaining errors are ours. All views expressed herein are our personal views, not necessarily aligned with those of the institution we are affiliated with.

Author contributions The two authors contributed equally to the manuscript.

Funding Olivier Gossner was financially supported by the French National Research Agency (ANR), "Investissements d'Avenir" (ANR-11-IDEX0003/LabEx Ecodec/ANR-11-LABX-0047) and by the European Union (SInfoNiA ERC- 2023-ADG 101142530).

Data availability Materials are publicly available and sources are indicated in the manuscript.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Arrow, Kenneth Joseph. 1965. *Aspects of the theory of risk-bearing*. Helsinki: Yrjö Jahnsonian Sätiö.

Beath, Alexander D. and Chris Flynn. 2018. Asset allocation, cost of investing and performance of European DB pension funds: The impact of real estate. CEM Benchmarking.

Beath, Alexander D. and Chris Flynn. 2020. Asset allocation and fund performance of defined benefit pension funds in the United States, 1998–2018. CEM Benchmarking.

Bermuda Monetary Authority. 2011. Insurance (prudential standards) (class 3A solvency requirement) rules 2011 BR 73 / 2011.

CFO-Forum. 2009. Market consistent embedded value basis for conclusions. http://www.cfoforum.eu.

Chetty, Raj. 2006. A new method of estimating risk aversion. *American Economic Review* 96 (5): 1821–1834.

Cox, John C., and Stephen A. Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3 (1–2): 145–166.

Cvitanić, Jakša, Huyen Pham, and Nizar Touzi. 1999. A closed-form solution to the problem of super-replication under transaction costs. *Finance and Stochastics* 3 (1): 35–54.

Czichowsky, Christoph, Rémi. Peyre, Walter Schachermayer, and Junjian Yang. 2018. Shadow prices, fractional Brownian motion, and portfolio optimisation under transaction costs. *Finance and Stochastics* 22 (1): 161–180.

Dana, Rose-Anne. 1995. An extension of Milleron, Mitjushin and Polterovich's result. *Journal of Mathematical Economics* 24 (3): 259–269.

Drèze, Jacques H. 1970. Market allocation under uncertainty. *European Economic Review* 2 (2): 133–165.

EIOPA. 2019a. Principles and Technical Specifications for the Common Framework: Annex 1 to Opinion on the practical implementation of the common framework for risk assessment and transparency for IORPs. BoS-19-246.

EIOPA. 2019b. Report on insurers' asset and liability management in relation to the illiquidity of their liabilities. BoS-19-593.

EIOPA. 2021. Consultation paper on the revision of the guidelines on the valuation of technical provisions. BoS-21-302.

EU Parliament. 2009. Directive 2009/138/EC of the European Parliament and of the council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II). Official Journal of the European Union, 52.

EU Parliament. 2015. Commission Delegated Regulation (EU) 2015/35. Official Journal of the European Union, 58.

FFA. 2021. L'assurance française: données clés 2020.

FINMA. 2020. Standardmodell Versicherungen Technische Beschreibung für das SST-Standardmodell Leben. https://www.finma.ch.

Fishburn, Peter C., and Burr R. Porter. 1976. Optimal portfolios with one safe and one risky asset: Effects of changes in rate of return and risk. *Management Science* 22 (10): 1064–1073.

GDV. 2020. Statistisches Taschenbuch der Versicherungswirtschaft 2020. https://www.gdv.de.

Gollier, Christian. 2001. *The economics of risk and time*. Cambridge: MIT press.

IASB. 2021. IFRS 17 Insurance Contracts. www.ifrs.org.

Jouini, Elyes, and Hédi. Kallal. 1995. Martingales and arbitrage in securities markets with transaction costs. *Journal of Economic Theory* 66 (1): 178–197.

Pratt, John W. 1964. Risk aversion in the small and in the large. *Econometrica* 122–136.

Ross, Stephen A. 1977. Risk, return, and arbitrage. In *Risk and return in finance*, vol. 1, ed. Irwin Friend and James L. Bicksler. Cambridge: American Telephone and TelegraphCompany., Ballinger Pub. Co.