Insurance: Mathematics and Economics 111 (2023) 41–56

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

# Multiple-prior valuation of cash flows subject to capital requirements

Hampus Engsner[a], Filip Lindskog[a,*], Julie Thøgersen[b]

[a] Stockholm University, Sweden
[b] Aarhus University, Denmark

ARTICLE INFO

*Article history:*
Received September 2021
Received in revised form February 2023
Accepted 24 February 2023
Available online 2 March 2023

*JEL classification:*
G22
G28
G32

*Keywords:*
Market-consistent valuation
Insurance valuation
Capital requirements
Limited liability
Optimal stopping

ABSTRACT

We study market-consistent valuation of liability cash flows motivated by current regulatory frameworks for the insurance industry. The value assigned to an insurance liability is the consequence of (1) considering a hypothetical transfer of an insurance company's liabilities, and financial assets intended to hedge these liabilities, to an empty corporate entity, and (2) considering the circumstances under which a capital provider would want to achieve and maintain ownership of this corporate entity given limited liability for the owner and that capital requirements have to be met at any time for continued ownership.

We focus on the consequences of the capital provider assessing the value of continued ownership in terms of a least favorable expectation of future dividends, meaning that we consider expectations with respect to probability measures in a set of equivalent martingale measures. We show that natural conditions on the set of probability measures imply that the value of a liability cash flow is given in terms of a solution to a backward recursion. Through a life and a non-life insurance example we demonstrate how to make the valuation approach operational.



## 1. Introduction

We consider the valuation of an aggregate insurance liability cash flow in run-off. The valuation approach is a direct consequence of considering a hypothetical transfer of the liability cash flow from an insurance company to an empty corporate entity set up with the sole purpose to manage the liability run-off. The owner of this entity needs to make sure at any time, in order to continue ownership of the entity, to pay claims and also to provide buffer capital according to the externally imposed solvency capital requirement (e.g. by a regulatory framework such as Solvency II).

The papers Möhr (2011) and Engsner et al. (2020, 2017) and the present paper approaches valuation of liabilities by considering a transfer of the liability to an empty corporate entity, called "reference undertaking" in the terminology of EIOPA (see European Commission, 2015, Article 38), and explores the consequence of the reference undertaking having to comply with repeated capital requirements throughout the run-off of the liability cash flow. A logical consequence of considering a capital provider/owner/shareholder, that one period at a time provides capital in order to satisfy the current capital requirement, is that the owner can decide not to provide further capital than what has already been provided. The owner who has made a net loss during the time of ownership can decide that enough is enough and simply give up the ownership. This means that the owner has limited liability.

The paper Engsner et al. (2017) considers only non-hedgeable liability cash flows and valuation as a consequence of an external capital provider providing capital given a sufficiently high expected one-period return on the capital in terms of cost-of-capital rates. The cost-of-capital approach in Engsner et al. (2017) builds on the approach to valuation presented in Möhr (2011) aiming to clarify the principles of valuation consistent with modern solvency regulatory frameworks.

The paper Engsner et al. (2020) considers a setting that takes the dependence of a liability cash flow on a financial market's price processes into account. In particular, Engsner et al. (2020) emphasizes that the liability cash flow, the replicating portfolio intended to hedge this cash flow, capital requirements, and a capital provider's risk aversion all together need to be specified in order to determine the value of a liability. The possibility of default is also considered in Hieber et al. (2019) that studies valuation of life-insurance contracts with guarantees in an incomplete-market setting with a single equivalent martingale measure used for valuation. However, capital requirements do not enter in the setup in Hieber et al. (2019) and therefore the valuation operator differs from that in Engsner et al. (2020).

* Corresponding author.
*E-mail addresses:* hampus.engsner@math.su.se (H. Engsner), lindskog@math.su.se (F. Lindskog), juliethoegersen@econ.au.dk (J. Thøgersen).

https://doi.org/10.1016/j.insmatheco.2023.02.007

The setting and valuation approach we consider are similar to those considered in Engsner et al. (2020). An essential difference is that here we consider a capital provider who assigns a value to possible future dividends and capital injections in terms of a least favorable expected value, where the expectations correspond to different equivalent martingale measures in an incomplete market setting. In a multiple-period setting, if we start with any such set $\mathcal{Q}$ of probability measures, then we may find that a probability measure that is the worst at a given time may not be the worst at a later time. We may think of cash flows from life insurance. If the capital provider at a given time financially benefits from survival of policyholders, then a conservative valuation from the risk-averse capital provider's perspective corresponds to a probability measure $\mathbb{Q}$ that assigns higher probability to the occurrences of deaths compared to the "real-world" probability measure $\mathbb{P}$. However, at a later time the composition of the portfolio may have changed so that the capital provider instead benefits from deaths of policyholders, and the $\mathbb{Q}$ considered earlier no longer corresponds to conservative valuation. In particular, we should consider optimization over sets $\mathcal{Q}$ of probability measures that are closed under pasting together a high-mortality probability measure with a low-mortality probability measure at stochastic times points. It turns out that this property of the set $\mathcal{Q}$ is not only a logical consequence of model uncertainty in multiple-period models but also necessary in order for the liability value to be computable via a backward recursion.

Insurance liability cash flows may be partly defined in terms of financial asset prices, specific interest rates or inflations indices. For liability cash flows where this is not the case, the cash flows may show significant correlation with market prices. Therefore, any insurance liability valuation methodology must be such not to introduce arbitrage opportunities and must consider replicating portfolios that hedge the financial component of a liability cash flow, whenever that is relevant. Consequently, there is a vast literature on market-consistent insurance valuation covering single-period, multiple-period and continuous-time valuation problems with varying assumptions on the financial market forming the basis for designing replicating portfolios of varying degrees of sophistication. We refer (in chronological order) to Grosen and Jørgensen (2002), Malamud et al. (2008), Wüthrich et al. (2011), Möhr (2011), Tsanakas et al. (2013), Wüthrich and Merz (2013), Pelsser and Stadje (2014), Engsner et al. (2017), Delong et al. (2019), Barigou and Dhaene (2019), Barigou et al. (2019), Engsner et al. (2020), and references therein.

A common theme in the literature on market-consistent insurance valuation is that the value assigned to a liability cash flow can be expressed as the sum of a market price of a replicating portfolio and a value assigned to the replication error (notice that a substantial replication error is a common feature of insurance liabilities). The liability values in this paper are also of this kind. Rebalancing times of a dynamic replicating portfolio means that the replication error has to be reassessed over time and taking this into consideration leads to the notion of time-consistent valuation. Similarly, repeated capital requirements lead to capital costs that are not known at the initial valuation time and taking such costs into account appropriately also require time-consistent valuation. Time consistency is a key concept in the literature on dynamic risk measurement. We refer (in chronological order) to Riedel (2004), Detlefsen and Scandolo (2005), Rosazza Gianin (2006), Cheridito et al. (2006), Artzner et al. (2007), Bion-Nadal (2008), Cheridito and Kupper (2009), Cheridito and Kupper (2011), and references therein.

In Artzner et al. (2020) and Deelstra et al. (2020) it is argued that diversifiable insurance risk should only be assigned a value corresponding to the $\mathbb{P}$-expectation of such risk since the law of large numbers applies if the insurance company may form arbitrarily large portfolios. In our setting this argument is not valid since the corporate entity to which the insurance company's aggregate liability is transferred is a separate entity that may not be merged with other corporate entities. In that sense the entity to which the liabilities are transferred may be seen as a special purpose vehicle. Although this entity benefits from diversification when capital requirements are computed, it can not diversify the liability further.

Optimal stopping with multiple priors for agents assessing risk in terms of dynamic convex risk measures is analyzed in Cheridito et al. (2006). Similar problems are analyzed in Engelage (2011), where the framework of optimal stopping with multiple priors in Riedel (2009) is extended to so-called dynamic variational preferences. From an applied perspective: whereas all priors/probability measures in a given set of priors are treated as equally likely in the framework in Riedel (2009), introducing (dynamic) penalty terms as in Cheridito et al. (2006) and Engelage (2011) means that the optimizing agent may assign different (dynamic) weights to the priors in the optimization problem. Solving optimal stopping problems with multiple priors numerically is a challenging computational problem. A general method for solving such problems is developed in Krätschmer et al. (2018).

Optimal stopping is a key element in our approach to valuation since the owner of the entity managing the run-off of the liability, just as shareholders in general, has limited liability. At any time, taking the value of assets and future liability cash flows into account, if a capital injection is needed to meet capital requirements, the owner may choose between making a capital injection or not. Without such a capital injection, ownership is terminated and the remaining assets are transferred to policyholders. Therefore, the rational owner determines optimal stopping times.

The approach we present for valuation of an insurance liability cash flow is the logical consequence of (1) considering a hypothetical transfer of an insurance company's liabilities, and financial assets intended to hedge these liabilities, to an empty corporate entity, and (2) considering the circumstances under which a capital provider would want to achieve and maintain ownership of this corporate entity given limited liability for the owner and that capital requirements have to be met at any time for continued ownership. We do not specify how the replicating portfolio, intended to hedge the liability cash flow, should be chosen. We do so because many possible static or dynamic hedging strategies may be considered and they all fit into the framework we present. We clarify under what fairly mild conditions the valuation operator is market consistent. In short, a sufficiently good hedging rule applied to the liability cash flow makes the valuation operator market consistent, whereas poor hedging does not.

The paper is organized as follows. Section 2 presents the valuation framework. Basic assumptions, notation and terminology are introduced in Subsection 2.1. Subsection 2.2 presents the valuation framework in the much simpler single-period setting in order to make key ideas and basic properties easily accessible. Subsection 2.3 presents the multiple-period valuation framework in full generality and presents a key result establishing the connection between the valuation problem formulated as a multiple-prior optimal stopping problem and as the solution to a backward recursion, with the valuation operator in the single-period setting appearing as a special case. Definitions and results are presented in Subsection 2.3 for general capital requirements. Subsection 2.4 then specializes by considering capital requirements given in terms of conditional monetary risk measures, in line with current regulatory frameworks. Conditions under which the valuation operator is market consistent are presented. These conditions are satisfied for a wide range of dynamic hedging strategies. Section 3 presents a general construction of parametric sets of probability measures that enable a wide range of insurance applications, and shows that the set of probability measures satisfies the properties making it suitable for optimal stopping with multiple priors. Section 4 considers a simple life insurance example that illustrates key features of the valuation framework and how the value is computed by solving a backward recursion. Section 5 considers a non-life insurance example for a model of the kind commonly used for claims reserving. The examples in Sections 4 and 5 show the need for the general theory developed in Section 3.

## 2. The valuation framework

### 2.1. Preliminaries

We consider time periods $1, \dots, T$, corresponding time points $0, 1, \dots, T$, and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)_{t=0}^T$ with $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \dots \subseteq \mathcal{F}_T = \mathcal{F}$, and $\mathbb{P}$ denotes the real-world measure. For $p \in [1, \infty)$, we write $L^p(\mathcal{F}_t, \mathbb{P})$ for the normed linear space of $\mathcal{F}_t$-measurable random variables $X$ with norm $\mathbb{E}^{\mathbb{P}}[|X|^p]^{1/p}$. We write $L^\infty(\mathcal{F}_t, \mathbb{P})$ for the normed linear space of $\mathcal{F}_t$-measurable essentially bounded random variables. We write $\mathbb{P}_t(\cdot)$ for $\mathbb{P}(\cdot \mid \mathcal{F}_t)$ and $\mathbb{E}_t^{\mathbb{P}}[\cdot]$ for $\mathbb{E}^{\mathbb{P}}[\cdot \mid \mathcal{F}_t]$. Equalities and inequalities between random variables should be interpreted in the $\mathbb{P}$-almost sure sense. A stopping time is a function $\tau : \Omega \to \{0, 1, \dots, T\} \cup \{+\infty\}$ such that $\{\tau = t\} \in \mathcal{F}_t$ for $t = 0, 1, \dots, T$.

For two probability measures $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)}$ equivalent to $\mathbb{P}$ and a stopping time $\tau \le T$, the probability measure $\mathbb{Q}^{(3)}(A) := \mathbb{E}^{\mathbb{Q}^{(1)}}[\mathbb{Q}^{(2)}(A \mid \mathcal{F}_\tau)]$, $A \in \mathcal{F}_T$, is called the pasting of $\mathbb{Q}^{(1)}$ and $\mathbb{Q}^{(2)}$ in $\tau$. It is often convenient to express the pasting $\mathbb{Q}^{(3)}$ of $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)}$ in $\tau$ in terms of the processes $D^{(1)}, D^{(2)}$ of Radon-Nikodym derivatives with respect to $\mathbb{P}$,

$$D_t^{(1)} = \frac{d\mathbb{Q}^{(1)}}{d\mathbb{P}}\Big|_{\mathcal{F}_t}, \quad D_t^{(2)} = \frac{d\mathbb{Q}^{(2)}}{d\mathbb{P}}\Big|_{\mathcal{F}_t}.$$

The pasting $\mathbb{Q}^{(3)}$ of $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)}$ in $\tau$ corresponds to

$$\begin{aligned} D_t^{(3)} &= \mathbb{I}\{t \le \tau\} D_t^{(1)} + \mathbb{I}\{t > \tau\} \frac{D_\tau^{(1)} D_t^{(2)}}{D_\tau^{(2)}} \\ &= \prod_{s=1}^t \left( \mathbb{I}\{s \le \tau\} \frac{D_s^{(1)}}{D_{s-1}^{(1)}} + \mathbb{I}\{s > \tau\} \frac{D_s^{(2)}}{D_{s-1}^{(2)}} \right). \end{aligned}$$

A set $\mathcal{Q}$ of probability measures equivalent to $\mathbb{P}$ is called stable under pasting if for any $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)} \in \mathcal{Q}$ and any stopping time $\tau \le T$, the pasting $\mathbb{Q}^{(3)}$ of $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)}$ in $\tau$ is an element in $\mathcal{Q}$. We call such a set stable under pasting. Such sets are also referred to as m-stable, time consistent or rectangular in the related literature.

We assume the existence of a financial market containing assets for which $\mathbb{F}$-adapted price processes $(S_t^0)_{t=0}^T$ and $(S_t^i)_{t=0}^T$, $i = 1, \dots, d$, are available. $(S_t^0)_{t=0}^T$ is the price process of a (predictable) locally riskless bond. These price processes generate the filtration $\mathbb{F}^S = (\mathcal{F}_t^S)_{t=0}^T \subset \mathbb{F}$. We assume that the financial market generated by the basic price processes and trading using $\mathbb{F}^S$-predictable trading strategies is arbitrage free and that any $\mathcal{F}_T^S$-measurable contingent claim is attainable. In particular, any $\mathcal{F}_T^S$-measurable contingent claim has a unique market price. We will also allow for $\mathbb{F}$-adapted cash flows that depend on insurance events and we will allow for $\mathbb{F}$-predictable trading strategies. With such cash flows added to the financial market we have an incomplete market setting.

We take the price process of the locally riskless bond as numéraire process and in what follows all financial values are discounted by this numéraire. This saves us from having to explicitly take interest rates processes into account and makes the mathematical expressions less involved.

We assume that the set $\mathcal{P}$ of equivalent martingale measures (for each $\mathbb{Q} \in \mathcal{P}$, $\mathbb{Q}$ is equivalent to $\mathbb{P}$ and the $(S_t^0)_{t=0}^T$-discounted price processes are $\mathbb{Q}$-martingales) is non-empty. By Proposition 6.43 in Föllmer and Schied (2016) the set $\mathcal{P}$ is stable under pasting. We will consider a non-empty subset $\mathcal{Q} \subset \mathcal{P}$. We refer to $\mathbb{Q} \in \mathcal{Q}$ as a market risk neutral probability measure. We emphasize that $\mathbb{E}^{\mathbb{Q}^{(1)}}[Z] = \mathbb{E}^{\mathbb{Q}^{(2)}}[Z]$ for any $\mathcal{F}_T^S$-measurable $Z$ and any $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)} \in \mathcal{Q}$.

We use the conventions $\sum_{l=k}^{k-1} := 0$ and $\inf \emptyset := +\infty$ for sums over an empty index set and the infimum of an empty set. We use the notation $(x)^+ := \max(0, x)$ and $x \wedge y := \min(x, y)$, and note that $r - (r - x)^+ = r \wedge x$.

### 2.2. Valuation in the single-period setting

In order to present the key ideas in an accessible way, without the mathematical details necessary for the general multiple-period setting presented in Section 2.3, we first focus on the single-period setting. We tacitly assume sufficient integrability of the relevant quantities appearing below.

We consider an insurance company with an aggregate insurance liability corresponding to a liability cash flow $X_1^o$ at the end of the period. Regulation forces the insurance company to comply with externally imposed capital requirements. The requirements put restrictions on the asset portfolio of the insurance company. A portfolio of traded assets is called a replicating portfolio and generates cash flow $X_1^r$ at the end of the period. The replicating portfolio is intended to, to some extent, offset the liability cash flow. $X_1 := X_1^o - X_1^r$ is the residual liability cash flow. We will, in accordance with current solvency regulation (Möhr (2011) and prescribed by EIOPA, see European Commission, 2015, Article 38) define the value of the liability cash flow $X^o$ by considering a hypothetical transfer, at the beginning of the period, of the liability and the replicating portfolio to a separate entity referred to as a reference undertaking. The reference undertaking has initially neither assets nor liabilities and its sole purpose is to manage the run-off of the liability. The benefit of ownership of the reference undertaking is the right to receive possible surplus at the end of the period. The capital requirement forces the reference undertaking to hold buffer capital $R_0$ in order for the transfer to occur. The amount $R_0$ should be financed jointly by the original insurance company and an agent aspiring ownership of the reference undertaking - the capital provider. The amount $C_0$ provided by the capital provider corresponds to the value the capital provider assigns at the beginning of the period to possible surplus at the end of the period. The remaining amount $V_0 := R_0 - C_0$ needs to be provided by the original insurance company. Given that the capital provider provides $C_0$ so that the transfer occurs, at the end of the period the policyholders receive

$$Z_{\text{ph}} = (R_0 + X_1^r) \wedge X_1^o = R_0 \wedge (X_1^o - X_1^r) + X_1^r,$$

i.e. the policyholders receive the amount $X_1^o$ they are promised if there is sufficient capital available, otherwise the available capital $R_0 + X_1^r$. The capital provider receives the surplus $Z_{\text{cp}} = (R_0 + X_1^r - X_1^o)^+$. Notice that the available capital at the end of the period $R_0 + X_1^r = Z_{\text{ph}} + Z_{\text{cp}}$ has an observable market price which we write $R_0 + \mathbb{E}^{\mathbb{Q}_0}[X_1^r]$. Suppose that the capital provider assigns the value

$$C_0 = \inf_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z_{\text{cp}}] = \inf_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[(R_0 + X_1^r - X_1^o)^+]$$

to the possible surplus at the end of the period, where $\mathcal{Q}$ is such that $\inf_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z]$ equals the market price $\mathbb{E}^{\mathbb{Q}_0}[Z]$ for any traded payoff $Z$. $C_0$ is a market-consistent value of $Z_{\text{cp}}$ in the sense that, for any traded payoff $Z$,

$$\inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z_{\text{cp}}+Z] = \inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z_{\text{cp}}] + \mathbb{E}^{\mathbb{Q}_0}[Z].$$

Consequently,

$$\begin{aligned} R_0 - C_0 + \mathbb{E}^{\mathbb{Q}_0}[X_1^r] &= \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z_{\text{ph}}] \\ &= \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[R_0 \wedge (X_1^o - X_1^r)] + \mathbb{E}^{\mathbb{Q}_0}[X_1^r] \end{aligned}$$

is a market-consistent value of $Z_{\text{ph}}$. Is it also a market-consistent value of the cash flow $X_1^o$ promised to the policyholders? The answer is "yes" if natural conditions are imposed. Suppose that the replication criterion used is such that if a traded payoff $Z$ is added to $X_1^o$, then the replicating portfolio hedging the liability cash flow $\widetilde{X}_1^o := X_1^o + Z$ generates the cash flow $\widetilde{X}_1^r = X_1^r + Z$. Suppose further that there is some (risk measure) $\rho$ with $\rho(0)=0$ such that $R_0 = \rho(X_1^r - X_1^o)$. Then

$$\begin{aligned} &\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[\rho(\widetilde{X}_1^r - \widetilde{X}_1^o) \wedge (\widetilde{X}_1^o - \widetilde{X}_1^r)] + \mathbb{E}^{\mathbb{Q}_0}[\widetilde{X}_1^r] \\ &= \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[\rho(X_1^r - X_1^o) \wedge (X_1^o - X_1^r)] + \mathbb{E}^{\mathbb{Q}_0}[X_1^r + Z] \end{aligned}$$

verifying that

$$V_0 + \mathbb{E}^{\mathbb{Q}_0}[X_1^r] = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[\rho(X_1^r - X_1^o) \wedge (X_1^o - X_1^r)] + \mathbb{E}^{\mathbb{Q}_0}[X_1^r]$$

is a market-consistent value of the liability cash flow $X_1^o$ (cf. Definition 3.1 in Pelsser and Stadje (2014)).

In Section 2.3 below we will consider the general multiple-period setting. That general setting is conceptually similar to the single-period setting but gives rise to various mathematical challenges. In particular, we need to pay careful attention to the choice of $\mathcal{Q}$ in the multiple-period setting. Theorem 1 below shows that, in the multiple-period setting, $V_0$ is given as the solution to a backward recursion:

$$V_t = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[R_t \wedge (X_{t+1}^o - X_{t+1}^r + V_{t+1})], \quad V_T = 0.$$

Market consistency of the value $V_0 + \sum_{t=1}^T \mathbb{E}^{\mathbb{Q}_0}[X_t^r]$ assigned to the liability cash flow $(X_t^o)_{t=1}^T$, that the original insurance company has promised its policyholders, holds under conditions similar to those presented above for the single-period setting, see Remark 3.

### 2.3. Valuation in the multiple-period setting

We consider an insurance company with an aggregate insurance liability corresponding to a liability cash flow given by the $\mathbb{F}$-adapted stochastic process $X^o = (X_t^o)_{t=1}^T$, and a replicating portfolio generating an $\mathbb{F}$-adapted cash flow $X^r = (X_t^r)_{t=1}^T$. Depending on the degree of replicability of the liability cash flow, the replicating portfolio could be anything from simply a position in the numéraire asset to a portfolio that is rebalanced dynamically according to some strategy. $X := X^o - X^r$ is the residual liability cash flow. Our approach to valuation is the logical consequence of considering a hypothetical transfer of the liability and the replicating portfolio to a separate entity referred to as a reference undertaking. The reference undertaking has initially neither assets nor liabilities and its sole purpose is to manage the run-off of the liability. The benefit of ownership is the right to receive certain dividends/surplus, defined below, until either the run-off of the liability cash flow is complete or until letting the reference undertaking default on its obligations to the policyholders. The term default means termination of ownership of the reference undertaking. The owner can be thought of as a shareholder with limited liability. Default therefore means exercise of the so-called limited-liability option. Regulation forces the owner of the reference undertaking to comply with externally imposed capital requirements at any time given continued ownership. The decision to default at time $t$ means to give up ownership and transfer the numéraire position $R_{t-1}$ and the remaining replicating portfolio cash flow $\sum_{s=t}^T X_s^r$ to the policyholders. The owner neither receives any dividend payment nor incurs any loss upon a decision to default, a consequence of limited liability. However, of course, it is possible that the owner has an accumulated loss at the time of default due to capital injected (in order to meet capital requirements) exceeding the surplus gains. A key feature of the multiple-period setting which did not enter in the single-period setting is that the owner of the reference undertaking needs to decide on a decision rule defining under which circumstances default occurs.

The default time is a stopping time $\tau \in \mathcal{S}_{1,T+1}$, where $\mathcal{S}_{t,T+1}$ denotes the set of $\mathbb{F}$ stopping times taking values in $\{t,\dots,T+1\}$. The event $\{\tau = T+1\}$ is to be interpreted as a complete liability run-off without default at any time. Formally,

$$\mathcal{S}_{t,T+1} := \{\tau : \tau \text{ is a stopping time with } \tau \geq t\} \wedge (T+1). \tag{1}$$

The cumulative cash flow to the owner can be written as

$$\sum_{t=1}^{\tau-1}(R_{t-1} - R_t - X_t), \quad X_t := X_t^o - X_t^r. \tag{2}$$

For ease of notation, define the payoff process $(H_t)_{t=1}^T$ by

$$H_1 := 0, \quad H_t := \sum_{s=1}^{t-1}(R_{s-1} - R_s - X_s) \quad \text{for } t > 1. \tag{3}$$

Note that this payoff process is predictable. The conservative value of the cash flow (2) is

$$\inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_0^{\mathbb{Q}}[H_\tau]. \tag{4}$$

We assume that the owner of the reference undertaking chooses a default time $\tau$ maximizing the value (4). Consequently, the value at time 0 of the reference undertaking is

$$\sup_{\tau\in\mathcal{S}_{1,T+1}} \inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_0^{\mathbb{Q}}[H_\tau]. \tag{5}$$

For $t \in \{1,\dots,T\}$, the value of the reference undertaking at time $t$, given no default at times $\leq t$, is given by the completely analogous expression upon replacing sup and inf in (5) by the essential supremum ess sup and essential infimum ess inf (see Appendix A.5 in Föllmer and Schied (2016) for details) and conditioning on $\mathcal{F}_t$ rather than $\mathcal{F}_0$. Notice that since no cash flows occur at times $> T$, the value of the reference undertaking is zero at time $T$. The value of the reference undertaking can thus be identified as the value of an American type derivative. Details on arbitrage-free pricing of American derivatives can be found in Section 6.3 in Föllmer and Schied (2016).

Since we are considering sets $\mathcal{Q}$ of probability measures we need the cash flows to be suitably integrable with respect to all $\mathbb{Q} \in \mathcal{Q}$. The following notion of uniform integrability, from Riedel (2009), will be used. The process $(H_t)_{t=1}^T$ in (3) is bounded by a $\mathcal{Q}$-uniformly integrable random variable in the sense that there exists $Z \geq 0$ such that

$$\sup_{t\in\{1,\dots,T\}} |H_t| \leq Z \text{ and } \lim_{K\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z\mathbb{I}_{\{Z\geq K\}}] = 0. \tag{6}$$

We now define the value of the reference undertaking, corresponding to what an external party would pay to become owner of the entity managing the run-off of the liability, and also the value of the residual liability. The sum of the latter and the market price of the replicating portfolio is the value of the original liability to policyholders and therefore is a theoretical aggregate premium.

**Definition 1.** Let $\mathcal{Q}$ be a set of market risk neutral probability measures. Consider sequences $(X_t)_{t=1}^T$ and $(R_t)_{t=0}^T$ with $X_t \in L^1(\mathcal{F}_t, \mathbb{Q})$ for $t \in \{1, \dots, T\}$ for every $\mathbb{Q} \in \mathcal{Q}$, $R_T = 0$ and $R_t \in L^1(\mathcal{F}_t, \mathbb{Q})$ for $t \in \{0, \dots, T-1\}$ for every $\mathbb{Q} \in \mathcal{Q}$. Define $C_T := 0$ and, for $t \in \{0, \dots, T-1\}$,

$$C_t := \operatorname*{ess\,sup}_{\tau \in \mathcal{S}_{t+1,T+1}} \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{s=t+1}^{\tau-1} (R_{s-1} - R_s - X_s)\Big], \tag{7}$$

where the set of stopping times $\mathcal{S}_{t+1,T+1}$ is given in (1). $C_t$ is *the value of the reference undertaking at time t* given no default at times $\le t$. $V_t := R_t - C_t$ is *the value of the residual liability at time t* given no default at times $\le t$.

Notice that

$$\begin{aligned} V_t &:= R_t - C_t \\ &= R_t - \operatorname*{ess\,sup}_{\tau \in \mathcal{S}_{t+1,T+1}} \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{s=t+1}^{\tau-1} (R_{s-1} - R_s - X_s)\Big] \\ &= \operatorname*{ess\,inf}_{\tau \in \mathcal{S}_{t+1,T+1}} \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[R_t - \sum_{s=t+1}^{\tau-1} (R_{s-1} - R_s - X_s)\Big] \\ &= \operatorname*{ess\,inf}_{\tau \in \mathcal{S}_{t+1,T+1}} \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{s=t+1}^{\tau-1} X_s + R_{\tau-1}\Big] \\ &\le \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{s=t+1}^{T} X_s\Big] =: \overline{V}_t. \end{aligned}$$

The general upper bound

$$\overline{V}_0 := \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_0^{\mathbb{Q}}\Big[\sum_{s=1}^{T} X_s\Big] \ge V_0 \tag{8}$$

does neither depend on the filtration nor on the capital requirements, and is typically much easier to compute than $V_0$. Therefore, this upper bound provides a useful conservative estimate of $V_0$. This statement is illustrated in the numerical example in Section 5. Notice that in general

$$\begin{aligned} V_t &= \operatorname*{ess\,inf}_{\tau \in \mathcal{S}_{t+1,T+1}} \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{s=t+1}^{\tau-1} X_s + R_{\tau-1}\Big] \\ &\ge \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}} \operatorname*{ess\,inf}_{\tau \in \mathcal{S}_{t+1,T+1}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{s=t+1}^{\tau-1} X_s + R_{\tau-1}\Big] =: \underline{V}_t. \end{aligned} \tag{9}$$

In particular, the general lower bound

$$\underline{V}_0 := \sup_{\mathbb{Q} \in \mathcal{Q}} \inf_{\tau \in \mathcal{S}_{1,T+1}} \mathbb{E}_0^{\mathbb{Q}}\Big[\sum_{s=1}^{\tau-1} X_s + R_{\tau-1}\Big] \le V_0 \tag{10}$$

may be attractive since it is typically easier to compute than $V_0$, see Section 5 for an illustration. Computing $\underline{V}_0$ means solving a standard optimal stopping problem for each $\mathbb{Q} \in \mathcal{Q}$ followed by finding the maximum of the obtained values $V_0^{\mathbb{Q}}$.

Notice that the value $L_0$ of the original liability cash flow $X^o$ follows directly from the procedure for transferring the liabilities and replicating portfolio to an external party (the new owner of the reference undertaking) accepting the transfer: $L_0$ equals the sum of the market value of the replicating portfolio and the value $V_0$ of the residual liability:

$$L_0 = \mathbb{E}_0^{\mathbb{Q}}\Big[\sum_{s=1}^{T} X_s^r\Big] + V_0,$$

where $\mathbb{Q}$ is any market risk neutral probability measure making the expectation equal the market value of the replicating portfolio. For details on the market consistency of the value $L_0$ we refer to Remark 3. Notice that the aggregate cash flow to the policyholders can be written

$$\mathbb{I}_{\{\tau=T+1\}} \sum_{t=1}^{T} X_t^o + \mathbb{I}_{\{\tau \le T\}} \Big(\sum_{t=1}^{\tau-1} X_t^o + R_{\tau-1} + \sum_{t=\tau}^{T} X_t^r\Big)$$

In particular, if $\mathcal{Q} = \{\mathbb{Q}\}$ and $\tau^*$ denotes an optimal default time (from the owner's perspective), then

$$\begin{aligned} L_0 &= \mathbb{E}^{\mathbb{Q}}\Big[\mathbb{I}_{\{\tau^*=T+1\}} \sum_{t=1}^{T} X_t^o + \mathbb{I}_{\{\tau^* \le T\}} \Big(\sum_{t=1}^{\tau^*-1} X_t^o \\ &\quad + R_{\tau^*-1} + \sum_{t=\tau^*}^{T} X_t^r\Big)\Big] \\ &= \mathbb{E}^{\mathbb{Q}}\Big[\mathbb{I}_{\{\tau^*=T+1\}} \sum_{t=1}^{T} X_t^o + \mathbb{I}_{\{\tau^* \le T\}} \Big(\sum_{t=1}^{\tau^*-1} X_t^o \\ &\quad + R_{\tau^*-1} + \mathbb{E}_{\tau^*}^{\mathbb{Q}}\Big[\sum_{t=\tau^*}^{T} X_t^r\Big]\Big)\Big]. \end{aligned}$$

Writing the liability value as the above expression demonstrates close connections to other valuation approaches such as Hieber et al. (2019) that also consider the possibility of default (cf. Theorem 3.3(a) in Hieber et al. (2019)).

We intend to build on the theory of multiple prior optimal stopping in Riedel (2009) where four assumptions on a set $\mathcal{Q}$ of probability measures are imposed in order for key results to hold. These assumptions are $\mathcal{Q}$-uniform integrability together with properties (i)-(iii) of the following definition.

**Definition 2.** A set $\mathcal{Q}$ of probability measures is suitable for multiple prior optimal stopping if the following properties hold. (i) Each $\mathbb{Q} \in \mathcal{Q}$ is equivalent to $\mathbb{P}$; (ii) $\mathcal{Q}$ is stable under pasting; (iii) For each $t \in \{0, \dots, T\}$,

$$\Big\{\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} : \mathbb{Q} \in \mathcal{Q}\Big\}$$

is weakly compact in $L^1(\mathcal{F}_T, \mathbb{P})$.

**Remark 1.** If $\mathcal{Q}$ satisfies the properties (i)-(iii) in Definition 2, then it follows from Theorem 2 in Riedel (2009) that the lower bound $\underline{V}_0$ in (10) equals $V_0$. This holds since for such $\mathcal{Q}$ the inequality in (9) is in fact a minimax identity. Notice also that for an arbitrary $\mathbb{Q}$ equivalent to $\mathbb{P}$, $\{\mathbb{Q}\}$ satisfies properties (i)-(iii) in Definition 2.

As a basis for applying the theory to be presented, we will later in Section 3 explicitly construct a useful set $\mathcal{Q}$ satisfying the properties in Definition 2 and present a detailed numerical example in Section 5.

We are now ready to state a key result which shows that $(C_t, V_t)$ defined in terms of a multiple prior optimal stopping problem may equivalently be defined as the solution to a backward recursion.

**Theorem 1.** *Let $\mathcal{Q}$ be a set of probability measures satisfying properties (i)-(iii) of Definition 2. Consider sequences $(X_t^o)_{t=1}^T$, $(X_t^r)_{t=1}^T$, $(R_t)_{t=0}^T$ with $X_t^o, X_t^r \in L^1(\mathcal{F}_t, \mathbb{Q})$ for $t \in \{1, \dots, T\}$ for every $\mathbb{Q} \in \mathcal{Q}$, $R_T = 0$ and $R_t \in L^1(\mathcal{F}_t, \mathbb{Q})$ for $t \in \{0, \dots, T-1\}$ for every $\mathbb{Q} \in \mathcal{Q}$. Set $X_t := X_t^o - X_t^r$ and assume that $(H_t)_{t=1}^T$ in (3) is bounded by a $\mathcal{Q}$-uniformly integrable random variable. Then*

*(i) If the sequences $(C_t)_{t=0}^T$ and $(V_t)_{t=0}^T$ are given by Definition 1, then*

$$C_t = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[(R_t - X_{t+1} - V_{t+1})^+], \quad C_T = 0, \tag{11}$$

$$V_t = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[R_t \wedge (X_{t+1} + V_{t+1})], \quad V_T = 0. \tag{12}$$

*(ii) The stopping times $(\tau_t^*)_{t=0}^{T-1}$ given by*

$$\tau_t^* = \inf\{s \in \{t+1, \dots, T\} : R_{s-1} - X_s - V_s < 0\} \wedge (T+1)$$

*are optimal in (7).*

*(iii) If the sequences $(C_t)_{t=0}^T$ and $(V_t)_{t=0}^T$ are given by (11) and (12), then, for $t \in \{0, \dots, T-1\}$, $C_t$ is given by (7) and $V_t = R_t - C_t$.*

We refer to Section 4 for a simple example illustrating how the backward recursion in Theorem 1 can be solved.

**Remark 2.** Stability under pasting of $\mathcal{Q}$ is a necessary requirement in Theorem 1. However, we show later in Theorem 6 that instead of the weak compactness property (iii) in Definition 2, which is assumed in Theorem 1, it is sufficient to verify weak relative compactness together with some natural additional properties. Notice that a bounded and uniformly integrable subset of $L^1(\mathcal{F}_t, \mathbb{P})$ is weakly relatively compact in $L^1(\mathcal{F}_t, \mathbb{P})$ (Theorem A.70 in Föllmer and Schied (2016)). Without weak compactness we can however not guarantee that there exists a $\mathbb{Q}^* \in \mathcal{Q}$ which solves the optimization problems (11) and (12).

**Proof of Theorem 1.** We will first consider the problem

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{S}_{t,T+1}} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[H_\tau]. \tag{13}$$

We define the multiple prior Snell envelope of $H$ with respect to $\mathcal{Q}$ as in Riedel (2009) by

$$U_{T+1}^{\mathcal{Q}} := H_{T+1}, \quad U_t^{\mathcal{Q}} := \max\Big\{H_t, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[U_{t+1}^{\mathcal{Q}}]\Big\} \quad \text{for } t \le T. \tag{14}$$

We know from Theorem 1 in Riedel (2009) that

$$U_t^{\mathcal{Q}} = \operatorname*{ess\,sup}_{\tau\in\mathcal{S}_{t,T+1}} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[H_\tau] \tag{15}$$

and that $\tau_t^* := \inf\{s \ge t : U_s^{\mathcal{Q}} = H_s\}$ is an optimal stopping time that solves (13). Define $\widetilde{U}^{\mathcal{Q}}$ by

$$\widetilde{U}_t^{\mathcal{Q}} := \operatorname*{ess\,sup}_{\tau\in\mathcal{S}_{t+1,T+1}} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[H_\tau].$$

We claim that the relation $\widetilde{U}_t^{\mathcal{Q}} = \operatorname{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[U_{t+1}^{\mathcal{Q}}]$ holds. Indeed, from (15),

$$U_t^{\mathcal{Q}} = \max\Big\{H_t, \operatorname*{ess\,sup}_{\tau\in\mathcal{S}_{t+1,T+1}} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[H_\tau]\Big\} = \max\big\{H_t, \widetilde{U}_t^{\mathcal{Q}}\big\}.$$

Therefore, from (14), we have the relation

$$\max\Big\{H_t, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[U_{t+1}^{\mathcal{Q}}]\Big\} = \max\big\{H_t, \widetilde{U}_t^{\mathcal{Q}}\big\}.$$

Since this holds for arbitrary adapted $H$, the claim is proved and gives

$$\begin{aligned}
C_t &= \widetilde{U}_t^{\mathcal{Q}} - H_{t+1} = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[U_{t+1}^{\mathcal{Q}}] - H_{t+1} \\
&= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[\max\{H_{t+1}, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{t+1}^{\mathbb{Q}}[U_{t+2}^{\mathcal{Q}}]\} - H_{t+1}] \\
&= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[\max\{0, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{t+1}^{\mathbb{Q}}[U_{t+2}^{\mathcal{Q}}] - H_{t+1}\}] \\
&= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[\max\{0, C_{t+1} + H_{t+2} - H_{t+1}\}] \\
&= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[(R_t - X_{t+1} - V_{t+1})^+].
\end{aligned}$$

Hence, we have shown (11) from which (12) is an immediate consequence. This concludes the proof of statement (i). □

### 2.4. Valuation with capital requirements by conditional monetary risk measures

We now consider the valuation problem in the setting where the capital requirements are given in terms of conditional monetary risk measures.

**Definition 3.** For $p \in [0, \infty]$ and $t \in \{0, \dots, T-1\}$, a conditional monetary risk measure is a mapping $\rho_t : L^p(\mathcal{F}_{t+1}, \mathbb{P}) \to L^p(\mathcal{F}_t, \mathbb{P})$ satisfying

$$\text{if } \lambda \in L^p(\mathcal{F}_t, \mathbb{P}) \text{ and } Y \in L^p(\mathcal{F}_{t+1}, \mathbb{P}),$$

$$\text{then } \rho_t(Y + \lambda) = \rho_t(Y) - \lambda, \tag{16}$$

$$\text{if } Y, \widetilde{Y} \in L^p(\mathcal{F}_{t+1}, \mathbb{P}) \text{ and } Y \le \widetilde{Y}, \text{ then } \rho_t(Y) \ge \rho_t(\widetilde{Y}), \tag{17}$$

$$\rho_t(0) = 0. \tag{18}$$

A sequence $(\rho_t)_{t=0}^{T-1}$ of conditional monetary risk measures is called a dynamic monetary risk measure.

The natural conditional monetary risk measures corresponding to current regulatory frameworks are defined in terms of conditional quantile functions. For integer $t \ge 0$, $x \in \mathbb{R}$, $u \in (0, 1)$ and $\mathcal{F}_{t+1}$-measurable $Z$, let

$$\begin{aligned}
F_{t,-Z}(x) &:= \mathbb{P}_t(-Z \le x), \\
F_{t,-Z}^{-1}(1-u) &:= \operatorname{ess\,inf}\{m \in L^0(\mathcal{F}_t, \mathbb{P}) : F_{t,-Z}(m) \ge 1-u\}
\end{aligned}$$

and define the conditional versions of value-at-risk and average value-at-risk as

$$\mathrm{V@R}_{t,u}(Z) := F_{t,-Z}^{-1}(1-u), \quad \mathrm{AV@R}_{t,u}(Z) := \frac{1}{u}\int_0^u \mathrm{V@R}_{t,v}(Z)dv.$$

Both $\mathrm{V@R}_{t,u}$ and $\mathrm{AV@R}_{t,u}$ are conditional monetary risk measures in the sense of Definition 3 for $p \ge 1$. Given conditional monetary risk measures $\rho_t : L^1(\mathcal{F}_{t+1}, \mathbb{P}) \to L^1(\mathcal{F}_t, \mathbb{P})$ we consider here

$$R_t := \rho_t(-X_{t+1} - V_{t+1}), \quad R_T := 0. \tag{19}$$

Notice that if $R_{t+1}$ is given and $C_{t+1}$ is given by Definition 1, then also $V_{t+1} := R_{t+1} - C_{t+1}$ is given and therefore $R_t$ is well defined by setting $R_t := \rho_t(-X_{t+1} - V_{t+1})$. Moreover, we may write

$$V_t := \varphi_t(X_{t+1} + V_{t+1}), \quad V_T := 0, \tag{20}$$

where $\varphi_t(Y) := \operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[\rho_t(-Y) \wedge Y]$.

**Theorem 2.** *Let $\rho_t : L^1(\mathcal{F}_{t+1}, \mathbb{P}) \to L^1(\mathcal{F}_t, \mathbb{P})$ be a conditional monetary risk measure in the sense of Definition 3 and let $\varphi_t : L^1(\mathcal{F}_{t+1}, \mathbb{P}) \to L^1(\mathcal{F}_t, \mathbb{P})$ be given by (20). Then*

$$\text{if } \lambda \in L^p(\mathcal{F}_t, \mathbb{P}) \text{ and } Y \in L^p(\mathcal{F}_{t+1}, \mathbb{P}),$$
$$\text{then } \varphi_t(Y + \lambda) = \varphi_t(Y) + \lambda, \tag{21}$$
$$\text{if } Y, \widetilde{Y} \in L^p(\mathcal{F}_{t+1}, \mathbb{P}) \text{ and } Y \le \widetilde{Y}, \text{ then } \varphi_t(Y) \le \varphi_t(\widetilde{Y}), \tag{22}$$
$$\varphi_t(0) = 0. \tag{23}$$

*Moreover, if $\widetilde{\rho}_t : L^1(\mathcal{F}_{t+1}, \mathbb{P}) \to L^1(\mathcal{F}_t, \mathbb{P})$ is a conditional monetary risk measure in the sense of Definition 3 such that $\rho_t \le \widetilde{\rho}_t$, then*

$$\varphi_t(Y) \le \widetilde{\varphi}_t(Y) := \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[\widetilde{\rho}_t(-Y) \wedge Y].$$

**Proof of Theorem 2.** The properties (21), (22) and (23) follow immediately by arguments similar to those in the proof of Proposition 1 in Engsner et al. (2017). The final property follows immediately since $x \wedge y$ is nondecreasing in $x$ and $y$. $\square$

Theorem 2 has consequences that should be seen as necessary requirements of any sound valuation method. If $X_1 + \cdots + X_T = c$ for some constant $c$, then the corresponding value $V_0 = c$. If we consider two residual liability cash flows $(X_t)_{t=1}^T$ and $(\widetilde{X}_t)_{t=1}^T$ such that $X_t \le \widetilde{X}_t$ for every $t$, then the corresponding values satisfy $V_0 \le \widetilde{V}_0$. Similarly, if the sequence of conditional monetary risk measures $(\rho_t)_{t=0}^{T-1}$ are replaced by a more prudent choice $(\widetilde{\rho}_t)_{t=0}^{T-1}$ such that $\rho_t \le \widetilde{\rho}_t$ for every $t$, then the corresponding values satisfy $V_0 \le \widetilde{V}_0$.

The following remark verifies market consistency of the valuation operator.

**Remark 3.** Suppose the conditions in Theorem 1 hold and that capital requirements $R_t$ are given in terms of conditional risk measures $\rho_t$ satisfying the conditions in Definition 3. Then the value of a liability cash flow $X^o$ to which a replicating portfolio generating the cash flow $X^r$ is assigned is given by

$$V_0 + \mathbb{E}^{\mathbb{Q}_0}\Big[\sum_{t=1}^T X_t^r\Big] = \varphi_0 \circ \cdots \circ \varphi_{T-1}\Big(\sum_{t=1}^T (X_t^o - X_t^r)\Big) + \mathbb{E}^{\mathbb{Q}_0}\Big[\sum_{t=1}^T X_t^r\Big], \tag{24}$$

where the second term on both sides of the equality sign is simply the market price of the replicating portfolio expressed in terms of an equivalent martingale measure. Note that $V_0$ can be expressed as a composition of mappings $\varphi_t$ due to the conditional cash additivity property (21). If the original liability cash flow $X^o$ is replaced by $\widetilde{X}^o = X^o + Z$, where the cash flow $\sum_{t=1}^T Z_t$ is fully replicable by hedging in the financial market, and if the replication criterion is such that $\sum_{t=1}^T \widetilde{X}_t^r = \sum_{t=1}^T X_t^r + \sum_{t=1}^T Z_t$, then it is clear that the expression in (24) defines a market-consistent valuation of the cash flow promised to the policyholders.

However, it should be emphasized that a regulator/financial supervisor may put restrictions on the complexity of replicating portfolios transferred along with the original liability of the insurance company. Consequently, market consistency of the value of the original liability cash flow given by (24) depends on whether the original insurance company is allowed sufficient flexibility in designing suitable hedging strategies. It should be emphasized that this issue is not specific to the approach to valuation considered here but applies to market consistency of liability valuations in general.

Let $(V_t^S)_{t=0}^T$ be given by $V_t^S = \sum_{u=1}^t X_u + V_t$, where $V_T = 0$,

$$V_t = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[R_t \wedge (X_{t+1} + V_{t+1})], \quad R_t = \rho_t(-X_{t+1} - V_{t+1}),$$

where $\rho_t$ is a suitable conditional monetary risk measure such as V@R$_t$ or AV@R$_t$. It is reasonable to require that $V^S$ is a $\mathbb{P}$-supermartingale which is equivalent to $V_t \ge \mathbb{E}_t^{\mathbb{P}}[X_{t+1} + V_{t+1}]$ which implies $V_t \ge \mathbb{E}_t^{\mathbb{P}}[X_{t+1} + \cdots + X_T]$. In particular, the $\mathbb{P}$-supermartingale property guarantees the existence of a nonnegative "risk margin" $V_t - \mathbb{E}_t^{\mathbb{P}}[X_{t+1} + \cdots + X_T] \ge 0$.

**Theorem 3.** *Let $X_t \in L^1(\mathcal{F}_t, \mathbb{P})$ for $t = 1, \dots, T$. Let $L^1(\mathcal{F}_{t+1}, \mathbb{P}) \ni Y_{t+1} \mapsto \rho_t(-Y_{t+1}) \in L^1(\mathcal{F}_t, \mathbb{P})$, for $t = 0, \dots, T-1$, be a conditional monetary risk measure such that*

$$\mathbb{E}_t^{\mathbb{P}}[\rho_t(-Y_{t+1}) - Y_{t+1}] > 0 \quad \text{or} \quad \mathbb{P}_t(\rho_t(-Y_{t+1}) - Y_{t+1} = 0) = 1. \tag{25}$$

*Then there exists a set $\mathcal{Q}$ of probability measures such that $(V_t^S)_{t=0}^T$ is a $\mathbb{P}$-supermartingale.*

**Proof of Theorem 3.** Notice that the supermartingale requirement is equivalent to

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[(R_t - X_{t+1} - V_{t+1})^+] \le \mathbb{E}_t^{\mathbb{P}}[R_t - X_{t+1} - V_{t+1}] \tag{26}$$

It is sufficient to find some $\mathbb{Q}$ such that the statement holds for $\mathcal{Q} = \{\mathbb{Q}\}$. We construct this $\mathbb{Q}$ by defining a suitable $\mathbb{P}$-martingale $(D_t)_{t=0}^T$ corresponding to the change of measure from $\mathbb{P}$ to $\mathbb{Q}$.

Let $W_{t+1} := R_t - X_{t+1} - V_{t+1}$, let $G_t(x) := \mathbb{P}_t(W_{t+1} \le x)$ denote the $\mathcal{F}_t$-conditional distribution function of $W_{t+1}$, and let $p_t := G_t(0)$. Let $(D_t)_{t=0}^T$, with $D_0 = 1$, be a $\mathbb{P}$-martingale satisfying

$$\frac{D_{t+1}}{D_t} = \begin{cases} 1 & \text{if } p_t \in \{0,1\}, \\ \exp\big(\lambda_t \Phi^{-1}(U_{t+1}) - \lambda_t^2/2\big) & \text{if } p_t \in (0,1), \end{cases}$$

where $U_{t+1}$ is independent of $\mathcal{F}_t$ and uniformly distributed on $(0,1)$ and, conditional of $\mathcal{F}_t$, $U_{t+1}$ and $W_{t+1}$ are countermonotone. Let $\lambda_t$ be some $\mathcal{F}_t$-measurable random variable satisfying

$$\exp\big(\lambda_t \Phi^{-1}(1-p_t) - \lambda_t^2/2\big)\mathbb{E}_t^{\mathbb{P}}\big[W_{t+1}^+\big] \le \mathbb{E}_t^{\mathbb{P}}\big[W_{t+1}\big] \quad \text{on } \{p_t \in (0,1)\}.$$

By construction,

$$\mathbb{E}_t^{\mathbb{P}}\Big[\frac{D_{t+1}}{D_t} W_{t+1}^+\Big] = \mathbb{E}_t^{\mathbb{P}}\big[W_{t+1}\big] \quad \text{on } \{p_t \in \{0,1\}\}.$$

Moreover, on $\{p_t \in (0,1)\}$,

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{P}}\Big[\frac{D_{t+1}}{D_t} W_{t+1}^+\Big] &= \int_0^{1-p_t} \exp\big(\lambda_t \Phi^{-1}(u) - \lambda_t^2/2\big) G_t^{-1}(1-u)du \\
&\le \exp\big(\lambda_t \Phi^{-1}(1-p_t) - \lambda_t^2/2\big) \int_0^{1-p_t} G_t^{-1}(1-u)du \\
&= \exp\big(\lambda_t \Phi^{-1}(1-p_t) - \lambda_t^2/2\big)\mathbb{E}_t^{\mathbb{P}}\big[W_{t+1}^+\big] \\
&\le \mathbb{E}_t^{\mathbb{P}}\big[W_{t+1}\big]. \quad \square
\end{aligned}$$

Property (25) in Theorem 3 is satisfied by AV@R$_{t,u}$ which is an example of so-called strictly expectation bounded risk measures, see Definition 5 and Example 3 in Rockafellar et al. (2006).

The following lemma is useful for constructing a bounding $\mathcal{Q}$-uniformly integrable random variable.

**Lemma 1.** *For any $\mathcal{Q}$-uniformly integrable $Z \geq 0$, $\operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]$ is a $\mathcal{Q}$-uniformly integrable random variable.*

**Proof of Lemma 1.** We need to show that

$$\lim_{K\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\Big[\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]\mathbb{I}_{\{\operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]\geq K\}}\Big]=0.$$

If we set $X = Z\mathbb{I}_{\{\operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]\geq K\}}$, then $X$ is $\mathcal{Q}$-uniformly integrable since it is of the form $Z\mathbb{I}_A$. Hence by the law of iterated expectations for $\mathcal{Q}$-uniformly integrable random variables (Lemma 1 in Riedel (2009)),

$$\begin{aligned}
&\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\Big[\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]\mathbb{I}_{\{\operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]\geq K\}}\Big]\\
&\quad= \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\Big[\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[X]\Big]\\
&\quad= \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[X].
\end{aligned}$$

Notice that $\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z] < \infty$ since for any $r > 0$,

$$\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z] \leq r + \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[|Z|\mathbb{I}_{|Z|>r}]$$

and, due to $\mathcal{Q}$-uniformly integrability of $Z$, we may choose $r$ to make the second term on the right-hand side sufficiently small. Since

$$\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\Big[\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z]\Big] = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[Z] < \infty,$$

the events $A_n = \{\operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Z] \geq n\}$ satisfy $\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{Q}(A_n) \to 0$ as $n\to\infty$. For any $\mathbb{Q}\in\mathcal{Q}$ and $r_n > 0$,

$$\sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\Big[Z\mathbb{I}_{A_n}\Big] \leq r_n \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{Q}(A_n) + \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\Big[|Z|\mathbb{I}_{\{|Z|>r_n\}}\Big]. \tag{27}$$

Consider a sequence $(r_n)_{n=1}^{\infty}$ such that $r_n \to \infty$ and $r_n \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{Q}(A_n) \to 0$ as $n\to\infty$. Applying this sequence to (27), taking the supremum over $\mathcal{Q}$ and letting $n\to\infty$ proves the statement of the lemma. □

The following theorem says that if the conditional monetary risk measures $\rho_t$ defining $R_t$ in (19) satisfy natural and verifiable bounds, then statements in Theorem 1 hold also in this setting.

**Theorem 4.** *Let $\mathcal{Q}$ be a set of probability measures satisfying properties (i)-(iii) of Definition 2. Consider sequences $(X_t^o)_{t=1}^T$, $(X_t^r)_{t=1}^T$, with $X_t^o, X_t^r \in L^1(\mathcal{F}_t,\mathbb{Q})$ for $t\in\{1,\dots,T\}$ for every $\mathbb{Q}\in\mathcal{Q}$, $R_T = 0$ and $R_t \in L^1(\mathcal{F}_t,\mathbb{Q})$ for $t\in\{0,\dots,T-1\}$ for every $\mathbb{Q}\in\mathcal{Q}$. Let $X_t := X_t^o - X_t^r$ and let $(R_t)_{t=0}^T$ be defined by (19). Assume that $\sum_{t=1}^T |X_t|$ is $\mathcal{Q}$-uniformly integrable. If the conditional monetary risk measures $\rho_t$ in (19) satisfy either*

$$|\rho_t(Z)| \leq K_\rho \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[|Z|] \text{ for some } K_\rho \in (1,\infty) \tag{28}$$

*or*

$$\mathbb{P}\in\mathcal{Q} \text{ and } |\rho_t(Z)| \leq K_\rho \mathbb{E}_t^{\mathbb{P}}[|Z|] \text{ for some } K_\rho \in (1,\infty), \tag{29}$$

*then $(H_t)_{t=0}^T$ defined in (3) satisfies that $\operatorname{ess\,sup}_{t=1,\dots,T} H_t$ is bounded by a $\mathcal{Q}$-uniformly integrable random variable. In particular, the statements in Theorem 1 hold.*

**Proof of Theorem 4.** Set

$$S_T := 0, \quad S_t := \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\Big[\sum_{u=t+1}^T |X_u|\Big] \text{ for } t = 0,1,\dots,T-1.$$

By Lemma 1 all variables $S_t$ are $\mathcal{Q}$-uniformly integrable. We will show by induction that, for all $t$, there exist constants $K_{V,t}, K_{R,t} \in (1,\infty)$ such that

$$|V_t| \leq K_{V,t} S_t, \quad |R_t| \leq K_{R,t} S_t \tag{30}$$

from which the statement of the theorem follows. Note that (30) trivially holds for $t = T$. In order to show the induction step, assume that (30) holds with $t$ replaced by $t+1$. If (29) holds, then

$$\begin{aligned}
|R_t| &= |\rho_t(-X_{t+1} - V_{t+1})|\\
&\leq K_\rho \mathbb{E}_t^{\mathbb{P}}[|X_{t+1}| + |V_{t+1}|]\\
&\leq K_\rho \mathbb{E}_t^{\mathbb{P}}[|X_{t+1}| + K_{t+1}^V S_{t+1}]\\
&\leq K_\rho K_{V,t+1} \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[|X_{t+1}| + S_{t+1}]\\
&= K_\rho K_{V,t+1} S_t,
\end{aligned}$$

where the law of iterated expectations for $\mathcal{Q}$-uniformly integrable random variables (Lemma 1 in Riedel (2009)) was used in the last step. If (28) holds, then similarly

$$\begin{aligned}
|R_t| &= |\rho_t(-X_{t+1} - V_{t+1})|\\
&\leq K_\rho \mathbb{E}_t^{\mathbb{Q}}[|X_{t+1}| + |V_{t+1}|]\\
&\leq K_\rho \mathbb{E}_t^{\mathbb{Q}}[|X_{t+1}| + K_{V,t+1} S_{t+1}]\\
&\leq K_\rho K_{V,t+1} \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[|X_{t+1}| + S_{t+1}]\\
&= K_\rho K_{V,t+1} S_t.
\end{aligned}$$

We also note that

$$\begin{aligned}
C_t &= \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[(R_t - X_{t+1} - V_{t+1})^+]\\
&\leq |R_t| + \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[|X_{t+1}| + |V_{t+1}|]\\
&\leq (K_\rho + 1) K_{V,t+1} S_t
\end{aligned}$$

which implies $|V_t| \leq |R_t| + C_t \leq (2K_\rho + 1)K_{V,t+1}S_t$. We have proved that (30) holds, i.e. the induction step. By the induction principle (30) holds for all $t$ and the proof is complete. □

## 3. Construction of sets of probability measures for multiple prior optimal stopping

Our aim here is to define a useful set $\mathcal{Q}$ of parametric probability measures that enables the analysis of a wide range of models and provides solutions to the multiple-prior optimization problem (7). In particular, the set $\mathcal{Q}$ constructed below will imply that optimization over $\mathcal{Q}$ can be reduced to optimization over the set of parameters, see Theorem 6 for the precise statement.

We will define a useful set of probability measures, satisfying all the requirements for applying key results on multiple prior optimal stopping, by defining the corresponding set of density processes $(D_{\lambda,t})_{t=0}^T$ of the form

$$D_{\lambda,0} := 1, \quad D_{\lambda,t} := \prod_{s=1}^t \int_\Theta f_s(\theta)\lambda_s(d\theta) \text{ for } t\in\{1,\dots,T\},$$

where $\Theta$ is a set of parameters and $(f_s)_{s=1}^T$ and $(\lambda_s)_{s=1}^T$ are defined below.

On $(\Omega, \mathcal{F}_T)$, a probability measure $\mathbb{Q}$ absolutely continuous with respect to $\mathbb{P}$ corresponds to a Radon-Nikodym derivative $D_T \in L^1(\mathcal{F}_T, \mathbb{P})$ and together with the filtration $(\mathcal{F}_t)_{t=0}^T$ give rise to the density process $(D_t)_{t=0}^T$ given by $D_t = \mathbb{E}_t^{\mathbb{P}}[D_T]$. Similarly, a set $\mathcal{Q}$ of probability measures, absolutely continuous with respect to $\mathbb{P}$, corresponds to the set $\mathcal{D}_T \subset L^1(\mathcal{F}_T, \mathbb{P})$ of Radon-Nikodym derivatives. Write $\overline{\mathcal{D}}_T$ for the $L^1$ closure of $\mathcal{D}_T$ and let $\overline{\mathcal{Q}}$ be the set of probability measures corresponding to the Radon-Nikodym derivatives $\overline{\mathcal{D}}_T$. For two probability measures $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)}$ with Radon-Nikodym derivatives $D_T^{(1)}, D_T^{(2)}$ the Radon-Nikodym derivative of the pasting of $\mathbb{Q}^{(1)}, \mathbb{Q}^{(2)}$ in $\tau$ is

$$D_\tau^{(1)} \frac{D_T^{(2)}}{D_\tau^{(2)}}.$$

The following result is both of independent interest and will be relevant for constructing stable sets of probability measures, depending on a parameter, that are useful for multiple prior optimal stopping problems.

**Theorem 5.** *Given $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbb{P})$, let $\mathcal{Q}$ be a set of probability measures equivalent to $\mathbb{P}$ that is convex and stable under pasting. Let $\mathcal{D}_T$ be the corresponding set of Radon-Nikodym derivatives and let $\overline{\mathcal{D}}_T$ be the $L^1$ closure of $\mathcal{D}_T$. Let $\mathcal{D}_t := \{D_t = \mathbb{E}_t^{\mathbb{P}}[D_T] : D_T \in \mathcal{D}_T\}$ and $\overline{\mathcal{D}}_t := \{D_t = \mathbb{E}_t^{\mathbb{P}}[D_T] : D_T \in \overline{\mathcal{D}}_T\}$. Then*

*(i) The set $\overline{\mathcal{Q}}$ corresponding to $\overline{\mathcal{D}}_T$ is convex and stable under pasting.*
*(ii) For each $t$, $\overline{\mathcal{D}}_t$ is convex and closed in $L^1(\mathcal{F}_T, \mathbb{P})$.*
*(iii) If $\mathcal{D}_T$ is $\mathbb{P}$-uniformly integrable, then for each $t$, $\mathcal{D}_t$ is weakly relatively compact in $L^1(\mathcal{F}_T, \mathbb{P})$ and $\overline{\mathcal{D}}_t$ is weakly compact in $L^1(\mathcal{F}_T, \mathbb{P})$.*
*(iv) If $\mathcal{D}_T$ is $\mathbb{P}$-uniformly integrable, then, for any $\mathcal{F}_{t+1}$-measurable $\mathcal{Q}$-uniformly integrable random variable $Y_{t+1}$,*

$$\operatorname*{ess\,inf}_{\mathbb{Q} \in \overline{\mathcal{Q}}} \mathbb{E}_t^{\mathbb{Q}}\big[Y_{t+1}\big] = \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}\big[Y_{t+1}\big]$$

*and similarly with* ess inf *replaced by* ess sup.

**Remark 4.** Since $\mathbb{E}^{\mathbb{P}}[D] = 1$ for $D \in \mathcal{D}_T$ (and similarly for $\overline{\mathcal{D}}_T$), by Lemma 4.10 in Kallenberg (2002), $\mathcal{D}_T$ is uniformly integrable if

$$\lim_{\mathbb{P}(A) \to 0} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{Q}(A) = 0$$

(and similarly for $\overline{\mathcal{D}}_T$).

**Proof of Theorem 5.** For both statements $(i)$ and $(ii)$, it is straightforward to verify that convexity holds so we only prove the remaining claims.

To prove $(i)$, consider any stopping time $\tau$ and $D_T^{(1)}, D_T^{(2)} \in \overline{\mathcal{D}}_T$. Take $(D_{n,T}^{(1)})_{n \geq 1}, (D_{n,T}^{(2)})_{n \geq 1} \subset \mathcal{D}_T$ such that $D_{n,T}^{(1)} \to D_T^{(1)}$ and $D_{n,T}^{(2)} \to D_T^{(2)}$ in $L^1$. Since $D_{n,T}^{(1)}, D_{n,T}^{(2)} \in \mathcal{D}_T$ and $\mathcal{Q}$ is stable under pasting, the Radon-Nikodym derivative of the pasting of $D_{n,T}^{(1)}, D_{n,T}^{(2)} \in \mathcal{D}_T$ in $\tau$ is also an element in $\mathcal{D}_T$. Therefore, statement (i) is proved if we show that there exists a subsequence $(n_i)$ such that

$$D_{n_i,\tau}^{(1)} \frac{D_{n_i,T}^{(2)}}{D_{n_i,\tau}^{(2)}} \to D_\tau^{(1)} \frac{D_T^{(2)}}{D_\tau^{(2)}} \text{ in } L^1. \tag{31}$$

Since, for $k = 1, 2$,

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}[|D_{n,T}^{(k)} - D_T^{(k)}|] &= \mathbb{E}^{\mathbb{P}}[\mathbb{E}_\tau^{\mathbb{P}}[|D_{n,T}^{(k)} - D_T^{(k)}|]] \\ &\geq \mathbb{E}^{\mathbb{P}}[|\mathbb{E}_\tau^{\mathbb{P}}[D_{n,T}^{(k)} - D_T^{(k)}]|] \\ &= \mathbb{E}^{\mathbb{P}}[|D_{n,\tau}^{(k)} - D_\tau^{(k)}|],\end{aligned}$$

we see that $D_{n,\tau}^{(k)} \to D_\tau^{(k)}$ in $L^1$. Since convergence in $L^1$ implies convergence in probability which in turn implies a.s. convergence along some subsequence $(n_i)$, we have

$$D_{n_i,\tau}^{(1)} \frac{D_{n_i,T}^{(2)}}{D_{n_i,\tau}^{(2)}} \to D_\tau^{(1)} \frac{D_T^{(2)}}{D_\tau^{(2)}} \text{ a.s.},$$

where we used the fact that $D_{n_i,\tau}^{(2)}$ and $D_\tau^{(2)}$ are strictly positive a.s. Since the terms of the sequence on the left-hand side are positive and all have expected values equal to 1, this sequence is uniformly integrable. Therefore, by Proposition 4.12 in Kallenberg (2002), the a.s. convergence can be replaced by convergence in $L^1$, i.e. (31) holds.

We now prove $(ii)$. Consider an $L^1$ convergent sequence $(D_{n,t})_{n \geq 1} \subset \overline{\mathcal{D}}_t$ with limit $D_t$. We will prove that $D_t \in \overline{\mathcal{D}}_t$. Take an arbitrary $D'_T \in \overline{\mathcal{D}}_T$ and let $D'_t := \mathbb{E}_t^{\mathbb{P}}[D'_T]$. Set

$$D_{n,T} := D_{n,t} \frac{D'_T}{D'_t}, \quad D_T := D_t \frac{D'_T}{D'_t}.$$

By construction $(D_{n,T})_{n \geq 1} \subset \overline{\mathcal{D}}_T$ and $D_{n,T} \to D_T$ in $L^1$. Hence, $D_T \in \overline{\mathcal{D}}_T$. Since $D_t = \mathbb{E}^{\mathbb{P}}[D_T] \in \overline{\mathcal{D}}_t$ and $D_{n,t} \to D_t$ in $L^1$ the proof of $(ii)$ is complete.

We now prove $(iii)$. From $(i)$ and $(ii)$ follow that $\overline{\mathcal{Q}}$ is convex and stable under pasting and, for each $t$, $\overline{\mathcal{D}}_t$ is a convex and closed subset of $L^1(\mathcal{F}_T, \mathbb{P})$. A convex and closed subset of $L^1(\mathcal{F}_T, \mathbb{P})$ is weakly closed (Theorem A.63 in Föllmer and Schied (2016)). A bounded and uniformly integrable subset of $L^1(\mathcal{F}_T, \mathbb{P})$ is weakly relatively compact (Theorem A.70 in Föllmer and Schied (2016)). Each $\mathcal{D}_t$ is a bounded subset of $L^1(\mathcal{F}_T, \mathbb{P})$:

$$\sup_{D_t \in \mathcal{D}_t} \|D_t\|_{L^1} = \sup_{D_t \in \mathcal{D}_t} \mathbb{E}^{\mathbb{P}}[D_t] = 1.$$

Hence, weak relative compactness of $\mathcal{D}_t$ follows if $\mathcal{D}_t$ is uniformly integrable. Similarly for $\overline{\mathcal{D}}_t$. Moreover, $\mathcal{D}_t$ is uniformly integrable if $\mathcal{D}_T$ is uniformly integrable, as the following argument shows. By Lemma 4.10 in Kallenberg (2002), since $\sup_{D \in \mathcal{D}_T} \mathbb{E}^{\mathbb{P}}[D] = 1$, $\mathcal{D}_T$ is $\mathbb{P}$-uniformly integrable if and only if $\lim_{\mathbb{P}(A) \to 0} \sup_{D \in \mathcal{D}_T} \mathbb{E}^{\mathbb{P}}[D; A] = 0$. If the latter holds, then in particular $\lim_{r \to \infty} \sup_{D \in \mathcal{D}_T} \mathbb{E}^{\mathbb{P}}[D; \mathbb{E}_t^{\mathbb{P}}[D] > r] = 0$ which is equivalent to

$$\lim_{r \to \infty} \sup_{D \in \mathcal{D}_T} \mathbb{E}^{\mathbb{P}}[\mathbb{E}_t^{\mathbb{P}}[D]; \mathbb{E}_t^{\mathbb{P}}[D] > r] = 0.$$

Hence, $\mathcal{D}_t$ is $\mathbb{P}$-uniformly integrable if $\mathcal{D}_T$ is $\mathbb{P}$-uniformly integrable. By the same argument, $\overline{\mathcal{D}}_t$ is uniformly integrable if $\overline{\mathcal{D}}_T$ is uniformly integrable. However, $\overline{\mathcal{D}}_T$ is uniformly integrable since it is the closure of a uniformly integrable set in $L^1$. Hence, $\overline{\mathcal{D}}_t$ is weakly compact in $L^1(\mathcal{F}_T, \mathbb{P})$ for every $t$. The proof of $(iii)$ is complete.

It remains to prove $(iv)$. Notice first that

$$\begin{aligned}\operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_t^{\mathbb{Q}}[Y_{t+1}] &= \operatorname*{ess\,inf}_{D \in \mathcal{D}_T} \frac{1}{D_t} \mathbb{E}_t^{\mathbb{P}}[D Y_{t+1}] \\ &= \operatorname*{ess\,inf}_{D \in \mathcal{D}_T} \frac{1}{D_t} \mathbb{E}_t^{\mathbb{P}}[D_{t+1} Y_{t+1}].\end{aligned}$$

Take $D^* \in \overline{\mathcal{D}}_T$ and $(D_n) \subset \mathcal{D}_T$ with $D_n \to D^*$ in $L^1$. Therefore, $D_n \xrightarrow{\mathbb{P}} D^*$ which implies $D_n Y_{t+1} \xrightarrow{\mathbb{P}} D^* Y_{t+1}$. Moreover, $D_{n,t} \xrightarrow{\mathbb{P}} D_t^*$ since $D_{n,t} = \mathbb{E}_t^{\mathbb{P}}[D_n]$, $D_t^* = \mathbb{E}_t^{\mathbb{P}}[D^*]$ and

$$\mathbb{E}^{\mathbb{P}}[|\mathbb{E}_t^{\mathbb{P}}[D_n] - \mathbb{E}_t^{\mathbb{P}}[D^*]|] \le \mathbb{E}^{\mathbb{P}}[\mathbb{E}_t^{\mathbb{P}}[|D_n - D^*|]] = \mathbb{E}^{\mathbb{P}}[|D_n - D^*|] \to 0$$

and convergence in $L^1$ implies convergence in probability. Since $Y_{t+1}$ is $\mathcal{Q}$-uniformly integrable, $\{DY_{t+1} : D \in \mathcal{D}_T\}$ is $\mathbb{P}$-uniformly integrable. Therefore, $D_nY_{t+1} \xrightarrow{\mathbb{P}} D^*Y_{t+1}$ implies $D_nY_{t+1} \to D^*Y_{t+1}$ in $L^1$. This further implies that $\mathbb{E}_t^{\mathbb{P}}[D_nY_{t+1}] \to \mathbb{E}_t^{\mathbb{P}}[D^*Y_{t+1}]$ in $L^1$ since

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}[|\mathbb{E}_t^{\mathbb{P}}[D_nY_{t+1}] - \mathbb{E}_t^{\mathbb{P}}[D^*Y_{t+1}]|]\\ &\quad\le \mathbb{E}^{\mathbb{P}}[\mathbb{E}_t^{\mathbb{P}}[|D_nY_{t+1} - D^*Y_{t+1}|]]\\ &\quad= \mathbb{E}^{\mathbb{P}}[|D_nY_{t+1} - D^*Y_{t+1}|] \to 0.\end{aligned}$$

In particular,

$$\frac{1}{D_{n,t}}\mathbb{E}_t^{\mathbb{P}}[D_nY_{t+1}] \xrightarrow{\mathbb{P}} \frac{1}{D_t^*}\mathbb{E}_t^{\mathbb{P}}[D^*Y_{t+1}]$$

which implies that there exists a subsequence $(n_i)$ such that

$$\frac{1}{D_{n_i,t}}\mathbb{E}_t^{\mathbb{P}}[D_{n_i}Y_{t+1}] \xrightarrow{\text{a.s.}} \frac{1}{D_t^*}\mathbb{E}_t^{\mathbb{P}}[D^*Y_{t+1}].$$

Therefore, for any $D^* \in \overline{\mathcal{D}}_T$,

$$\frac{1}{D_t^*}\mathbb{E}_t^{\mathbb{P}}[D^*Y_{t+1}] \ge \operatorname*{ess\,inf}_{D\in\mathcal{D}_T}\frac{1}{D_t}\mathbb{E}_t^{\mathbb{P}}[DY_{t+1}].$$

The same argument shows the corresponding identity for ess sup. The proof is complete. □

Consider a parameter set $\Theta$ which is taken to be a subset of a complete and separable metric space. For each $t \in \{1,\dots,T\}$, let $f_t \ge 0$ be a measurable function on $\Omega \times \Theta$ such that $\omega \mapsto f_t(\omega,\theta)$ is $\mathcal{F}_t$-measurable for each $\theta \in \Theta$. It is assumed that the $\mathcal{F}_t$-measurable random variables $f_t(\theta)$ satisfy

$$\operatorname*{ess\,inf}_{\theta\in\Theta} f_t(\theta) > 0 \ \mathbb{P}\text{-a.s. for } t \in \{1,\dots,T\}, \tag{32}$$

$$\mathbb{E}_{t-1}^{\mathbb{P}}[f_t(\theta)] = 1 \text{ for } (t,\theta) \in \{1,\dots,T\}\times\Theta. \tag{33}$$

For each $t \in \{1,\dots,T\}$, let $\lambda_t$ be an $\mathcal{F}_{t-1}$-measurable random element in the space $\mathcal{P}(\Theta)$ of probability measures on $\Theta$ equipped with the topology of weak convergence. Let $\Lambda_t$ be the set of all such random probability measures. For $\lambda_t \in \Lambda_t$ for all $t$, let

$$D_{\lambda,T} := \prod_{t=1}^{T}\int_{\Theta} f_t(\theta)\lambda_t(\mathrm{d}\theta), \quad D_{\lambda,t} := \mathbb{E}_t^{\mathbb{P}}[D_{\lambda,T}] \text{ for } t < T. \tag{34}$$

Notice that, due to properties (32) and (33), $(D_{\lambda,t})_{t=0}^T$ is a positive $\mathbb{P}$-martingale with $\mathbb{E}^{\mathbb{P}}[D_{\lambda,t}] = 1$. Let

$$\mathcal{D}_{f,T} := \left\{\prod_{t=1}^{T} f_t(\theta) : \theta \in \Theta\right\}, \quad \widetilde{\mathcal{D}}_{f,T} := \left\{D_{\lambda,T} : \lambda_t \in \Lambda_t \text{ for all } t\right\}$$

and let $\overline{\mathcal{D}}_{f,T}$ be the $L^1$-closure of $\mathcal{D}_{f,T}$. For $t = 0,1,\dots,T-1$, let

$$\mathcal{D}_{f,t} := \left\{D_t = \mathbb{E}_t^{\mathbb{P}}[D_T] : D_T \in \mathcal{D}_{f,T}\right\}$$

and let $\widetilde{\mathcal{D}}_{f,t}$ and $\overline{\mathcal{D}}_{f,t}$ be defined analogously.

**Definition 4.** Denote by $\mathcal{Q}_\Theta, \widetilde{\mathcal{Q}}_\Theta, \overline{\mathcal{Q}}_\Theta$ the sets of probability measures corresponding to the sets $\mathcal{D}_{f,T}, \widetilde{\mathcal{D}}_{f,T}, \overline{\mathcal{D}}_{f,T}$ of Radon-Nikodym derivatives with respect to $\mathbb{P}$.

Notice that $\mathcal{Q}_\Theta$ corresponds to only considering measures $\lambda_t(\cdot) = 1_{\{\theta\}}(\cdot)$ in (34). We will show in Theorem 6 that $\overline{\mathcal{Q}}_\Theta$ has the properties assumed in Theorem 1. We also show that Theorem 1 holds also for $\widetilde{\mathcal{Q}}_\Theta$.

**Theorem 6.** *Consider the sets $\mathcal{Q}_\Theta, \widetilde{\mathcal{Q}}_\Theta, \overline{\mathcal{Q}}_\Theta$ and $\mathcal{D}_{f,T}, \widetilde{\mathcal{D}}_{f,T}, \overline{\mathcal{D}}_{f,T}$ in Definition 4.*

*(i) The sets $\widetilde{\mathcal{Q}}_\Theta$ and $\overline{\mathcal{Q}}_\Theta$ are convex and stable under pasting.*
*(ii) For every $t \in \{1,\dots,T\}$, $\overline{\mathcal{D}}_{f,t}$ is closed in $L^1(\mathcal{F}_T,\mathbb{P})$.*
*(iii) If $\widetilde{\mathcal{D}}_{f,T}$ is $\mathbb{P}$-uniformly integrable, then $\widetilde{\mathcal{D}}_{f,t}$ is weakly relatively compact in $L^1(\mathcal{F}_T,\mathbb{P})$ and $\overline{\mathcal{D}}_{f,t}$ is weakly compact in $L^1(\mathcal{F}_T,\mathbb{P})$ for every $t \in \{0,\dots,T\}$.*
*(iv) If $\widetilde{\mathcal{D}}_{f,T}$ is $\mathbb{P}$-uniformly integrable, then, for any $\mathcal{F}_{t+1}$-measurable $\widetilde{\mathcal{Q}}_\Theta$-uniformly integrable random variable $Y_{t+1}$,*

$$\begin{aligned}\operatorname*{ess\,inf}_{\mathbb{Q}\in\overline{\mathcal{Q}}_\Theta}\mathbb{E}_t^{\mathbb{Q}}[Y_{t+1}] &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\widetilde{\mathcal{Q}}_\Theta}\mathbb{E}_t^{\mathbb{Q}}[Y_{t+1}]\\ &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{Q}_\Theta}\mathbb{E}_t^{\mathbb{Q}}[Y_{t+1}]\\ &= \operatorname*{ess\,inf}_{\theta\in\Theta}\mathbb{E}_t^{\mathbb{P}}[Y_{t+1}f_{t+1}(\theta)]\end{aligned}$$

*and similarly with* ess inf *replaced by* ess sup.

**Proof of Theorem 6.** We first prove $(i)$. We first prove convexity and stability under pasting for the set of probability measures with Radon-Nikodym derivatives $D_T \in \widetilde{\mathcal{D}}_{f,T}$ with respect to $\mathbb{P}$. We first prove convexity. Note that for a density process $(D_t)_{t=0}^T$ with $D_T \in \widetilde{\mathcal{D}}_{f,T}$,

$$\frac{D_{t+1}}{D_t} = \frac{\prod_{s=1}^{t+1}\int_\Theta f_s(\theta)\lambda_s(\mathrm{d}\theta)}{\prod_{s=1}^{t}\int_\Theta f_s(\theta)\lambda_s(\mathrm{d}\theta)} = \int_\Theta f_{t+1}(\theta)\lambda_{t+1}(\mathrm{d}\theta).$$

Consider density processes $D^{(1)}, D^{(2)}$ with $D_T^{(1)}, D_T^{(2)} \in \widetilde{\mathcal{D}}_{f,T}$, let $c \in (0,1)$ and set $D^{(3)} := cD^{(1)} + (1-c)D^{(2)}$. Then

$$\begin{aligned}\frac{D_{t+1}^{(3)}}{D_t^{(3)}} &= \frac{1}{D_t^{(3)}}\Big(cD_t^{(1)}\int_\Theta f_{t+1}(\theta)\lambda_{t+1}^{(1)}(\mathrm{d}\theta)\\ &\quad + (1-c)D_t^{(2)}\int_\Theta f_{t+1}(\theta)\lambda_{t+1}^{(2)}(\mathrm{d}\theta)\Big)\\ &= \int_\Theta f_{t+1}(\theta)\left(\frac{1}{D_t^{(3)}}\Big(cD_t^{(1)}\lambda_{t+1}^{(1)} + (1-c)D_t^{(2)}\lambda_{t+1}^{(2)}\Big)(\mathrm{d}\theta)\right)\end{aligned}$$

and the convexity property follows since

$$\frac{1}{D_t^{(3)}}\Big(cD_t^{(1)}\lambda_{t+1}^{(1)} + (1-c)D_t^{(2)}\lambda_{t+1}^{(2)}\Big) \in \Lambda_{t+1}.$$

We now prove stability under pasting. Consider density processes $D^{(1)}, D^{(2)}$ with $D_T^{(1)}, D_T^{(2)} \in \widetilde{\mathcal{D}}_{f,T}$, and let $\tau$ be a stopping time. Then

$$\begin{aligned}D_t^{(3)} &:= \prod_{s=1}^{t}\left(\mathbb{I}\{s\le\tau\}\frac{D_s^{(1)}}{D_{s-1}^{(1)}} + \mathbb{I}\{s>\tau\}\frac{D_s^{(2)}}{D_{s-1}^{(2)}}\right)\\ &= \prod_{s=1}^{t}\int_\Theta f_s(\theta)\Big(\mathbb{I}\{s\le\tau\}\lambda_s^{(1)} + \mathbb{I}\{s>\tau\}\lambda_s^{(2)}\Big)(\mathrm{d}\theta)\end{aligned}$$

and since $\mathbb{I}\{s\le\tau\}, \mathbb{I}\{s>\tau\}$ are $\mathcal{F}_{s-1}$-measurable,

$$\mathbb{I}\{s\le\tau\}\lambda_s^{(1)} + \mathbb{I}\{s>\tau\}\lambda_s^{(2)} \in \Lambda_s,$$

which proves stability under pasting. Theorem 5(*i*) completes the proof of (*i*).

Notice that (*ii*) follows immediately from (*i*) together with Theorem 5(*ii*). Similarly, (*iii*) follows immediately from (*i*) together with Theorem 5(*iii*).

It remains to prove (*iv*). The two last identities in (*iv*) follow from the definition of $\mathcal{Q}_\Theta$ and $\widetilde{\mathcal{Q}}_\Theta$:

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\widetilde{\mathcal{Q}}_\Theta}\mathbb{E}_t^{\mathbb{Q}}\big[Y_{t+1}\big]=\operatorname*{ess\,inf}_{\lambda_{t+1}\in\Lambda_{t+1}}\mathbb{E}_t^{\mathbb{P}}\left[Y_{t+1}\int_\Theta f_{t+1}(\theta)\lambda_{t+1}(\mathrm{d}\theta)\right]$$
$$=\operatorname*{ess\,inf}_{\theta\in\Theta}\mathbb{E}_t^{\mathbb{P}}\big[Y_{t+1}f_{t+1}(\theta)\big].$$

The first identity follows from Theorem 5(*iv*). The proof is complete. □

As noted in Riedel (2009), if the sample space $\Omega$ is finite, weak compactness follows if the set of priors is weakly closed. This allows us to find a simpler set of measures than $\overline{\mathcal{Q}}_\Theta$ in Theorem 6.

**Theorem 7.** *Consider a filtered probability space $(\Omega,(\mathcal{F}_t)_{t=0}^T,\mathbb{P})$, where $\Omega$ is finite. Let $\Theta$ be a compact set. For each $t=1,\dots,T$, let $f_t\ge 0$ be a measurable function on $\Omega\times\Theta$ such that $\omega\mapsto f_t(\omega,\theta)$ is $\mathcal{F}_t$-measurable for each $\theta\in\Theta$, such that $\theta\mapsto f_t(\omega,\theta)$ is continuous for each $\omega\in\Omega$, and such that (32) and (33) hold. Then the set*

$$\mathcal{D}^\alpha_{f,T}:=\left\{\prod_{t=1}^T f_t(\alpha_t):(\alpha_t)_{t=1}^T\text{ is }\Theta\text{-valued and predictable}\right\}$$

*of Radon-Nikodym derivatives is uniformly bounded and weakly compact in $L^1(\mathcal{F}_T,\mathbb{P})$. The set of measures $\mathcal{Q}^\alpha_\Theta$ corresponding to $\mathcal{D}^\alpha_{f,T}$ is stable under pasting.*

**Proof.** Write $\Omega=\{\omega_1,\dots,\omega_n\}$. Without loss of generality, assume that $\mathbb{P}(\omega_k)>0$ for all $k$. Consider any sequence $(\alpha^{(i)})_{i=1}^\infty$ of predictable processes $\alpha^{(i)}=(\alpha_t^{(i)})_{t=1}^T$. We may write $\alpha^{(i)}:\Omega\to\Theta^T$. Since $\Theta^T$ is a compact set and $\Omega$ is finite, there exists some subsequence $\alpha^{(n_i)}$ and some $\alpha:\Omega\to\Theta^T$ such that $\alpha^{(n_i)}(\omega_k)\to\alpha(\omega_k)$ for each $k=1,\dots,n$. $\alpha$ is also predictable since $\alpha^{(n_i)}\to\alpha$ a.s. Hence, by the continuity of $(f_t)_{t=1}^T$, for any sequence of the form $(\prod_{t=1}^T f_t(\alpha_t^{(i)}))_i$, there exists an almost surely convergent subsequence $\prod_{t=1}^T f_t(\alpha_t^{(n_i)})\to\prod_{t=1}^T f_t(\alpha_t)$. This ensures almost sure closure of $\mathcal{D}^\alpha_{f,T}$, which in turn implies weak closure. Hence the set is weakly compact. Uniform boundedness follows since, for each of the finitely many pairs $(t,\omega_k)$, $f_t(\omega_k,\cdot)$ is a continuous function on a compact set and therefore bounded. Hence, there exists $K\in(0,\infty)$ such that $D_T<K$ for all $D_T\in\mathcal{D}^\alpha_{f,T}$. It remains to verify stability under pasting. Consider $D_T^{(1)},D_T^{(2)}\in\mathcal{D}^\alpha_{f,T}$, defined by the processes $\alpha^{(1)},\alpha^{(2)}$. Then, for any stopping time $\tau$, the process $\alpha^{(3)}:=(\mathbb{I}_{\{\tau>t\}}\alpha_t^{(1)}+\mathbb{I}_{\{\tau\le t\}}\alpha_t^{(2)})_{t=1}^T$ is a predictable process, and $D_T^{(3)}\in\mathcal{D}^\alpha_{f,T}$ defined by $\alpha^{(3)}$ is the pasting of $D_T^{(1)}$ and $D_T^{(2)}$ in $\tau$. □

**Remark 5.** We now explain how the objects $f_t(\theta)$ relate to parametric models for stochastic processes. Consider a $d$-dimensional stochastic process $(X_t)_{t=1}^T$, and let $\mathcal{F}_t=\sigma(X_1,\dots,X_t)$. Let

$$F_t(x):=\mathbb{P}(X_t\le x\mid\mathcal{F}_{t-1})\quad\text{and}\quad F_t(x,\theta):=\mathbb{Q}_\theta(X_t\le x\mid\mathcal{F}_{t-1}),$$
$$\theta\in\Theta.$$

Assume that $\mathbb{P}$ and $\mathbb{Q}_\theta$ are equivalent for all $\theta\in\Theta$. Let $f_t(\theta):=\frac{\mathrm{d}F_t(\cdot,\theta)}{\mathrm{d}F_t}(X_t)$, where $\frac{\mathrm{d}G}{\mathrm{d}H}$ is interpreted as the Radon-Nikodym derivative (if it exists) between two distribution functions. Then

$$\frac{\mathrm{d}\mathbb{Q}_\theta}{\mathrm{d}\mathbb{P}}=\prod_{t=1}^T f_t(\theta).$$

If $\Omega=\{\omega_1,\dots,\omega_n\}$, then the situation simplifies. Let $\{x_1,\dots x_n\}$ be the set in which $X_t$ takes values. If we define the probability mass functions

$$p_t(x_i):=\mathbb{P}(X_t=x_i\mid\mathcal{F}_{t-1})\ \text{ and }\ p_t(x_i,\theta):=\mathbb{Q}_\theta(X_t=x_i\mid\mathcal{F}_{t-1}),$$

then $f_t(\theta)=\frac{p_t(X_t,\theta)}{p_t(X_t)}$. Therefore, in order to use Theorem 7 we need only ensure equivalence between the measures for all parameters under consideration and to verify that $\theta\mapsto p_t(x_i,\theta)$ is continuous on $\Theta$. As we will see in Section 4, we need not necessarily concern ourselves with the exact form of $f_t(\theta)$.

**Remark 6.** It might be useful to discuss the numerical procedure to calculate $V_t$ in a Markovian setting. Consider a $d$-dimensional Markov chain $(X_t)_{t=1}^T$, i.e. a $d$-dimensional stochastic process satisfying the Markov property

$$\mathbb{P}(X_{t+1}\le x\mid X_t,\dots,X_1)=\mathbb{P}(X_{t+1}\le x\mid X_t)=:F(x,X_t,t).$$

Let the filtration be generated by $(X_t)_{t=1}^T$, $\mathcal{F}_t=\sigma(X_1,\dots,X_t)$ and, for simplicity, let the residual cash flow be given by $(X_{t,1})_{t=1}^T$, i.e. the first component of $(X_t)_{t=1}^T$. Assume also that $(X_t)_{t=1}^T$ has the Markov property with respect to all $\mathbb{Q}_\theta\in\mathcal{Q}_\Theta$:

$$\mathbb{Q}_\theta(X_{t+1}\le x\mid X_t,\dots,X_1)=\mathbb{Q}_\theta(X_{t+1}\le x\mid X_t)=:F(x,X_t,t,\theta).$$

Assume that the conditions in Theorem 6 hold, and choose $\overline{\mathcal{Q}}_\Theta$ as our set of priors. By Theorem 6, we have that

$$\operatorname*{ess\,sup}_{\mathbb{Q}\in\overline{\mathcal{Q}}_\Theta}\mathbb{E}_t^{\mathbb{Q}}\big[g(X_{t+1})\big]=\operatorname*{ess\,sup}_{\theta\in\Theta}\mathbb{E}^{\mathbb{Q}_\theta}\big[g(X_{t+1})\big]$$
$$=\operatorname*{ess\,sup}_{\theta\in\Theta}\int g(x)\mathrm{d}F(x,X_t,t,\theta)$$

for any $\overline{\mathcal{Q}}_\Theta$-uniformly integrable random variable of the form $g(X_{t+1})$, where $g:\mathbb{R}^d\to\mathbb{R}$. It can be shown by induction that if the risk measure $\rho_t$ depends only on the $\mathcal{F}_t$-conditional law of the input, then $V_t$ is some function of $X_t$. We write $V_t=:v_t(X_t)$. The setting in the example in Section 4 satisfies the above assumptions, with state space determined by $(N_t)_{t=0}^2$. The simplest case to explore is the one where the sample space is finite. Assume $X_t$ can take values in $\{x_1,\dots x_n\}$. The recursion is then performed backwards in the following way. At time $t$, assume we have calculated $v_{t+1}(x_i)$ for each $i=1,\dots,n$. Note that $V_T=0$, so we start this procedure at $T-1$. We then calculate $v_t(x_i)$ for each $i=1,\dots,n$ by performing the following calculations: We first calculate $R_t^{(i)}=\rho_t(-X_{t+1,1}-V_{t+1})=\rho_t(-X_{t+1,1}-v_{t+1}(X_{t+1}))$ with respect to the law of $X_{t+1,1}+v_{t+1}(X_{t+1})$ conditional on $X_t=x_i$. We then solve the optimization problem

$$v_t(x_i)=\operatorname*{ess\,sup}_{\theta\in\Theta}\mathbb{E}^{\mathbb{Q}_\theta}[R_t^{(i)}\wedge(X_{t+1,1}+v_{t+1}(X_{t+1}))\mid X_t=x_i]\tag{35}$$

We continue with this recursion backward in time until we reach $t=0$. The potential problems with this method are twofold. Firstly, solving (35) may be nontrivial. Secondly, even if the state space is finite, it tends to grow exponentially with the dimension $d$ so the number of required calculations may become too large. To address the second problem, one may need to make approximations of the functions $v_t$.

## 4. Life insurance example

Consider a simple life insurance contract issued for a cohort of size $N_0$ at time 0. Let the process $(N_t)_{t=0}^2$ describe the evolution of the cohort size at times 0, 1, 2. Suppose that each contract pays out 2 monetary units at the end of each year the policyholder is alive and 3 upon the death of the policyholder. The aggregate liability cash flow $(X_1^o, X_2^o)$ is fully non-hedgeable so $(X_1^r, X_2^r) = (0, 0)$, and $(X_1, X_2) = (X_1^o, X_2^o)$ is given by

$$X_1 = 2N_1 + 3(N_0 - N_1) = 3N_0 - N_1,$$
$$X_2 = 2N_2 + 3(N_1 - N_2) = 3N_1 - N_2.$$

We assume that the probability of death for a member of the cohort is $p \in (0, 1)$ for each of the two years, and that death events are independent. This means that $\mathcal{L}^{\mathbb{P}}(N_t \mid N_{t-1}) = \text{Bin}(N_{t-1}, 1-p)$ for $t = 1, 2$, where the superscript $\mathbb{P}$ means that the conditional distribution is with respect to $\mathbb{P}$.

For the valuation procedure, we chose $\rho_t = \text{V@R}_{t,1-\beta}$. Since $(N_t)_{t=0}^2$ takes values in a finite set, $(N_t)_{t=0}^2$ can be defined on a finite filtered probability space with filtration generated by $(N_t)_{t=0}^2$. Consider probability measures $\mathbb{Q}_\theta$ such that $\mathcal{L}^{\mathbb{Q}_\theta}(N_t \mid N_{t-1}) = \text{Bin}(N_{t-1}, 1-\theta)$, for $t = 1, 2$, and consider the set $\mathcal{Q}_{[\theta^l,\theta^u]} = \{\mathbb{Q}_\theta : \theta \in [\theta^l, \theta^u]\}$, where $0 < \theta^l \le \theta^u < 1$. A natural interpretation could be that $[\theta^l, \theta^u]$ represents an uncertainty region for the probability of death within the cohort, while $p$ represents a point estimate of the same probability. Equivalence between measures in $\mathcal{Q}_{[\theta^l,\theta^u]}$ and $\mathbb{P}$ is obvious. Furthermore, note that the probability mass function of $\text{Bin}(n, \theta)$ is continuous in the parameter $\theta$. However, $\mathcal{Q}_{[\theta^l,\theta^u]}$ is not stable under pasting. As shown in Theorem 7, a weakly compact set of probability measures that is stable under pasting is constructed as follows. Consider a predictable process $\alpha = (\alpha_t)_{t=1}^2$ taking values in $[\theta^l, \theta^u]$. Then define $\mathbb{Q}_\alpha$ by the following conditional laws of $(N_t)_{t=0}^2$:

$$\mathcal{L}^{\mathbb{Q}_\alpha}(N_t \mid \mathcal{F}_{t-1}) = \text{Bin}(N_{t-1}, \alpha_t).$$

The set $\mathcal{Q}_\alpha$ of all such measures $\mathbb{Q}_\alpha$ is stable under pasting and weakly compact. Intuitively, instead of having a constant mortality probability, we have a variable one, selected randomly at the beginning of each year.

We now compute the residual liability value $V_0$ for the choice $\mathcal{Q} = \mathcal{Q}_\alpha$. Denote by $q_\beta(n, r)$ the $\beta$-quantile of $\text{Bin}(n, r)$. Using the fact that $\mathcal{L}^{\mathbb{P}}(N_1 - N_2 \mid \mathcal{F}_1) = \text{Bin}(N_1, p)$,

$$\begin{aligned}(X_2 + V_2) \wedge R_1 &= (3N_1 - N_2) \wedge R_1\\ &= (3N_1 - N_2) \wedge \text{V@R}_{1,1-\beta}(N_2 - 3N_1)\\ &= 2N_1 + (N_1 - N_2) \wedge \text{V@R}_{1,1-\beta}(N_2 - N_1)\\ &= 2N_1 + (N_1 - N_2) \wedge q_\beta(N_1, p).\end{aligned}$$

Note that $\mathcal{L}^{\mathbb{Q}_\theta}(N_1 - N_2 \mid \mathcal{F}_1) = \text{Bin}(N_1, \theta)$, and $V_1$ is given by

$$\begin{aligned}V_1 &= \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}_\alpha} \mathbb{E}_1^{\mathbb{Q}}[(X_2 + V_2) \wedge R_1]\\ &= \operatorname*{ess\,sup}_{\theta \in [\theta^l,\theta^u]} \mathbb{E}_1^{\mathbb{Q}_\theta}[(3N_1 - N_2) \wedge R_1]\\ &= 2N_1 + \operatorname*{ess\,sup}_{\theta \in [\theta^l,\theta^u]} \sum_{n=0}^{N_1} \binom{N_1}{n} (1-\theta)^{N_1-n} \theta^n (n \wedge q_\beta(N_1, p)) \qquad (36)\\ &= 2N_1 + \sum_{n=0}^{N_1} \binom{N_1}{n} (1-\theta^u)^{N_1-n} (\theta^u)^n (n \wedge q_\beta(N_1, p)).\end{aligned}$$

The last equality follows from the fact that (36) is an increasing function of $\theta$ for any value of $N_1$. Note that

$$\begin{aligned}X_1 + V_1 &= 3N_0 + N_1\\ &\quad + \sum_{n=0}^{N_1} \binom{N_1}{n} (1-\theta^u)^{N_1-n} (\theta^u)^n (n \wedge q_\beta(N_1, p))\\ &=: y_1(N_1),\end{aligned}$$

is increasing in $N_1$. Completely analogously to the derivation of $V_1$,

$$\begin{aligned}V_0 &= \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}_\alpha} \mathbb{E}_0^{\mathbb{Q}}[(X_1 + V_1) \wedge R_0] \qquad (37)\\ &= \operatorname*{ess\,sup}_{\theta \in [\theta^l,\theta^u]} \sum_{n=0}^{N_0} \binom{N_0}{n} \theta^{N_0-n} (1-\theta)^n (y_1(n) \wedge y_1(q_\beta(N_0, 1-p)))\\ &= \sum_{n=0}^{N_0} \binom{N_0}{n} (\theta^l)^{N_0-n} (1-\theta^l)^n (y_1(n) \wedge y_1(q_\beta(N_0, 1-p))).\end{aligned}$$

Hence, the measure $\mathbb{Q}$ that solves the optimization problem in (37) corresponds to the (deterministic) process $(\alpha_t)_{t=1}^2$ given by $\alpha_1 = \theta^l$ and $\alpha_2 = \theta^u$.

## 5. Non-life insurance example

In this section we consider an application illustrating the general theory presented up to this point, in particular we illustrate the need for the construction in Section 3. The model we consider is of the kind commonly used for non-life claims reserving with data in the form of claims triangles. The model is compatible with the classical chain-ladder method, see e.g. Mack (1993).

We consider a setting where the liability cash flow is Gaussian and independent of traded asset prices, both under $\mathbb{P}$ and under any $\mathbb{Q}_\theta \in \mathcal{Q}_\Theta$. We consider two cases:

Case 1 In this case we assume that $\mathcal{Q} = \mathcal{Q}_\Theta$. Recall that $\mathcal{Q}_\Theta$ is not stable under pasting: this decision maker does not consider probability measures corresponding to switching between probability measures in $\mathcal{Q}_\Theta$ depending on information revealed over time.

Case 2 In this case we assume that $\mathcal{Q} = \widetilde{\mathcal{Q}}_\Theta$. This decision maker exhibits a behavior that is time consistent. Notice that $\widetilde{\mathcal{Q}}_\Theta$ is considerably larger than $\mathcal{Q}_\Theta \subset \widetilde{\mathcal{Q}}_\Theta$.

The liability cash flow is assumed to be fully non-hedgeable by financial assets and consequently we take $X^r = 0$ which means that $X = X^o$. In order to make the illustration clear, we choose $T = 2$. Let $C_{i,k} := C_{i,k}^{\text{orig}}/v_i$ denote the exposure adjusted cumulative amount of payments to policyholders for accident year $i$, where $v_i$ is a known exposure measure for accident year $i$. The evolution of the exposure adjusted cumulative amounts is assumed to satisfy

$$C_{i,1} = \beta_0^{\mathbb{P}} + \frac{\sigma_0^{\mathbb{P}}}{\sqrt{v_i}} \varepsilon_{i,1}, \quad C_{i,2} = \beta_1^{\mathbb{P}} C_{i,1} + \frac{\sigma_1^{\mathbb{P}}}{\sqrt{v_i}} \varepsilon_{i,2},$$

where all $\varepsilon_{i,k}$ are independent and $N(0, 1)$ with respect to $\mathbb{P}$. Suppose that we are uncertain about the parameter values and want to consider probability measures $\mathbb{Q}_\theta$, $\theta = (\beta_0, \sigma_0, \beta_1, \sigma_1)$, such that

$$C_{i,1} = \beta_0 + \frac{\sigma_0}{\sqrt{v_i}} \varepsilon_{i,1}^\theta, \quad C_{i,2} = \beta_1 C_{i,1} + \frac{\sigma_1}{\sqrt{v_i}} \varepsilon_{i,2}^\theta,$$

where all $\varepsilon_{i,k}^\theta$ are independent and $N(0, 1)$ with respect to $\mathbb{Q}_\theta$. We choose a parameter set $\Theta \subset (0, \infty)^4$ that describes the uncertainty about the parameter values.

Suppose that $C_{i,k}$ with $i+k \le 0$ are observed at time 0 and that $C_{i,k}$ with $i+k=t$, $t=1,2$, are observed at times $t>0$ and therefore contain cash flows that are part of the outstanding liability to the policyholders. The (incremental) liability cash flow $X=(X_1,X_2)$ is given by

$$X=\Big(v_{-1}(C_{-1,2}-C_{-1,1})+v_0C_{0,1},\, v_0(C_{0,2}-C_{0,1})\Big).$$

Notice that $C_{-1,1}$ is here considered to be a known constant. Direct computations give

$$\mathbb{E}^{\mathbb{P}}[X_1+X_2]=v_{-1}(\beta_1^{\mathbb{P}}-1)C_{-1,1}+v_0\beta_0^{\mathbb{P}}\beta_1^{\mathbb{P}},$$
$$\mathbb{E}^{\mathbb{Q}_\theta}[X_1+X_2]=v_{-1}(\beta_1-1)C_{-1,1}+v_0\beta_0\beta_1.$$

The filtration is given by the $\sigma$-algebras $\mathcal{F}_0=\{\emptyset,\Omega\}$, $\mathcal{F}_1=\sigma(\varepsilon_{-1,2},\varepsilon_{0,1})$ and $\mathcal{F}_2=\sigma(\varepsilon_{0,2})\vee\mathcal{F}_1$. In order to have the correct evolution with respect to $\mathbb{Q}_\theta$ of the cumulative amounts it is seen that we must require that

$$\mathbb{Q}_\theta(\varepsilon_{-1,2}\in\cdot\mid\mathcal{F}_0)\sim N(\mu_{-1,2},\sigma^2_{-1,2}),$$
$$\mu_{-1,2}=\frac{\beta_1-\beta_1^{\mathbb{P}}}{\sigma_1^{\mathbb{P}}/\sqrt{v_{-1}}}C_{-1,1},\ \sigma_{-1,2}=\frac{\sigma_1}{\sigma_1^{\mathbb{P}}},$$
$$\mathbb{Q}_\theta(\varepsilon_{0,1}\in\cdot\mid\mathcal{F}_0)\sim N(\mu_{0,1},\sigma^2_{0,1}),$$
$$\mu_{0,1}=\frac{\beta_0-\beta_0^{\mathbb{P}}}{\sigma_0^{\mathbb{P}}/\sqrt{v_0}},\ \sigma_{0,1}=\frac{\sigma_0}{\sigma_0^{\mathbb{P}}},$$
$$\mathbb{Q}_\theta(\varepsilon_{0,2}\in\cdot\mid\mathcal{F}_1)\sim N(\mu_{0,2},\sigma^2_{0,2}),$$
$$\mu_{0,2}=\frac{\beta_1-\beta_1^{\mathbb{P}}}{\sigma_1^{\mathbb{P}}/\sqrt{v_0}}C_{0,1},\ \sigma_{0,2}=\frac{\sigma_1}{\sigma_1^{\mathbb{P}}}.$$

This corresponds to, in the setting of Section 3, choosing

$$f_1(\theta)=\frac{\varphi(\varepsilon_{-1,2};\mu_{-1,2},\sigma^2_{-1,2})\varphi(\varepsilon_{0,1};\mu_{0,1},\sigma^2_{0,1})}{\varphi(\varepsilon_{-1,2};0,1)\varphi(\varepsilon_{0,1};0,1)}, \tag{38}$$
$$f_2(\theta)=\frac{\varphi(\varepsilon_{0,2};\mu_{0,2},\sigma^2_{0,2})}{\varphi(\varepsilon_{0,2};0,1)},$$

where $\varphi(x;\mu,\sigma^2)$ denotes the density function of $N(\mu,\sigma^2)$. By Remark 4, the set $\widetilde{\mathcal{D}}_{f,2}$ is $\mathbb{P}$-uniformly integrable if

$$\lim_{\mathbb{P}(A)\to 0}\sup_{\mathbb{Q}\in\widetilde{\mathcal{Q}}_\Theta}\mathbb{Q}(A)=0$$

which holds here since $\sigma_k^{\mathbb{P}}/\sigma_k$ and $|\beta_k-\beta_k^{\mathbb{P}}|$ both take values in bounded intervals bounded away from 0. The sets $A\in\mathcal{F}_2$ are of type $\{(\varepsilon_{-1,2},\varepsilon_{0,1},\varepsilon_{0,2})\in B\}$ for measurable sets $B\subset\mathbb{R}^3$ such that $\mathbb{P}((\varepsilon_{-1,2},\varepsilon_{0,1},\varepsilon_{0,2})\in B)\to 0$. Therefore, it follows from Theorem 6 that the set $\widetilde{\mathcal{Q}}_\Theta$ in Definition 4 satisfies the requirements for multiple prior optimal stopping. In particular, Theorem 1 holds with $\mathcal{Q}=\widetilde{\mathcal{Q}}_\Theta$.

$\Theta$ can be chosen to reflect parameter uncertainty. To illustrate how such a choice may be implemented, consider the regression estimators from Lindholm et al. (2017) based on data from accident years $i=i_0,\dots,-1$:

$$\widehat{\beta_0^{\mathbb{P}}}=\frac{\sum_{i=i_0}^{-1}v_iC_{i,1}}{\sum_{i=i_0}^{-1}v_i},\quad \widehat{(\sigma_0^{\mathbb{P}})^2}=\frac{1}{-i_0-1}\sum_{i=i_0}^{-1}v_i(C_{i,1}-\widehat{\beta_0^{\mathbb{P}}})^2,$$
$$\widehat{\beta_1^{\mathbb{P}}}=\frac{\sum_{i=i_0}^{-2}v_iC_{i,1}C_{i,2}}{\sum_{i=i_0}^{-2}v_iC_{i,1}^2},$$
$$\widehat{(\sigma_1^{\mathbb{P}})^2}=\frac{1}{-i_0-2}\sum_{i=i_0}^{-2}v_i(C_{i,2}-\widehat{\beta_1^{\mathbb{P}}}C_{i,1})^2.$$

Here $i_0$ denotes the index of the first accident year observed. These estimators are unbiased and uncorrelated. We now proceed with a numerical illustration, with parameter values $(\beta_0^{\mathbb{P}},\sigma_0^{\mathbb{P}},\beta_1^{\mathbb{P}},\sigma_1^{\mathbb{P}})=(2/3,1/5,3/2,1/5)$, $i_0=-10$, and $v_i=1$ for $i=-10,\dots,0$. Based on these parameter values and a large number $n$ of simulated independent standard normal $\varepsilon_{i,j}$, leading to $n$ iid copies of $C_{-10,1},\dots,C_{-1,1},C_{-10,2},\dots,C_{-2,2}$, we estimate $(\beta_0^{\mathbb{P}},\sigma_0^{\mathbb{P}},\beta_1^{\mathbb{P}},\sigma_1^{\mathbb{P}})$ $n$ times. Fig. 1 presents scatter plots, which suggests that the iid vectors of estimators are approximately $N_4(\mu,\Sigma)$-distributed, where $\mu$ and $\Sigma$ are the sample mean and sample covariance matrix. We can therefore shape an approximative confidence region with confidence level $p$ of the parameter values by the squared Mahalanobis distance as

$$\Theta=\Big\{z\in\mathbb{R}^4:(z-\mu)^{\mathrm{T}}\Sigma^{-1}(z-\mu)\le F^{-1}_{\chi^2(4)}(p)\Big\}$$
$$=\Big\{\mu+rLs\in\mathbb{R}^4:r^2\le F^{-1}_{\chi^2(4)}(p),\ s\in\mathbb{S}^3\Big\},$$

where $L$ is the (lower triangular) Cholesky decomposition of $LL^{\mathrm{T}}=\Sigma$, $F_{\chi^2(4)}$ is the distribution function of the $\chi^2(4)$ and $\mathbb{S}^3$ is the unit sphere in $\mathbb{R}^4$. For the evaluation at time 1, only $(\beta_1,\sigma_1)$ needs to be considered, leading to a set $\Theta_{\beta_1,\sigma_1}\subset\mathbb{R}^2$ satisfying that

$$\{(0,0,\beta_1,\sigma_1):(\beta_1,\sigma_1)\in\Theta_{\beta_1,\sigma_1}\}$$

is the orthogonal projection of $\Theta$ onto the $(\beta_1,\sigma_1)$ coordinate plane: $\beta_0=\sigma_0=0$. Explicitly,

$$\Theta_{\beta_1,\sigma_1}=\Big\{z\in\mathbb{R}^2:(z-\mu_{\beta_1,\sigma_1})^{\mathrm{T}}\Sigma^{-1}_{\beta_1,\sigma_1}(z-\mu_{\beta_1,\sigma_1})\le F^{-1}_{\chi^2(4)}(p)\Big\}$$
$$=\Big\{\mu_{\beta_1,\sigma_1}+rL_{\beta_1,\sigma_1}s\in\mathbb{R}^2:r^2\le F^{-1}_{\chi^2(4)}(p),\ s\in\mathbb{S}^1\Big\},$$

where $\mathbb{S}^1$ is the unit sphere in $\mathbb{R}^2$, $\mu_{\beta_1,\sigma_1}$ is the subvector of the last two entries of $\mu$ and $L_{\beta_1,\sigma_1}$ is the Cholesky decomposition of the submatrix $\Sigma_{\beta_1,\sigma_1}$ of $\Sigma$. Similarly, to compute the upper bound of $V_0$ in (39), only $(\beta_0,\beta_1)$ need to be considered, leading to a similar set $\Theta_{\beta_0,\beta_1}\subset\mathbb{R}^2$.

The left plot in Fig. 1 shows a scatted plot of 1000 iid estimates of $(\beta_0^{\mathbb{P}},\beta_1^{\mathbb{P}})$ together with boundaries $\partial\Theta_{\beta_0,\beta_1}$ for $p=0.1$ (blue) and for $p=0.9$ (red). The right plot in Fig. 1 shows a scatted plot of 1000 iid estimates of $(\beta_1^{\mathbb{P}},\sigma_1^{\mathbb{P}})$ together with boundaries $\partial\Theta_{\beta_1,\sigma_1}$ for $p=0.1$ (blue) and for $p=0.9$ (red).

Let $\rho_0,\rho_1$ be conditional monetary risk measures defined in terms of conditional quantiles with respect to $\mathbb{P}$, such as, for $t=0,1$, $\rho_t=\mathrm{V@R}_{t,p}$ or $\rho_t=\mathrm{AV@R}_{t,p}$. In both cases, $c:=\rho_0(e_1^{\mathbb{P}})=\rho_1(e_2^{\mathbb{P}})$ is a constant for an $\mathcal{F}_{t+1}$-measurable $e_t^{\mathbb{P}}\sim N(0,1)$ and independent of $\mathcal{F}_t$ with respect to $\mathbb{P}$. Then

$$R_1=\rho_1(-X_2)=\rho_1(-\mathbb{E}_1^{\mathbb{P}}[X_2]+\mathrm{Var}_1^{\mathbb{P}}(X_2)^{1/2}e_2^{\mathbb{P}})$$
$$=\mathbb{E}_1^{\mathbb{P}}[X_2]+\mathrm{Var}_1^{\mathbb{P}}(X_2)^{1/2}c$$
$$=v_0(\beta_1^{\mathbb{P}}-1)C_{0,1}+\sqrt{v_0}\sigma_1^{\mathbb{P}}c.$$

### 5.1. Case 1: computing upper and lower bounds for $V_0$

In this case $\mathcal{Q}=\mathcal{Q}_\Theta$ does not satisfy the conditions of Theorem 1 and therefore we can not compute $V_0$ by backward recursion. However, upper and lower bounds for $V_0$ are easily computed. From (8) we have the upper bound

$$V_0\le\sup_{\mathbb{Q}\in\mathcal{Q}_\Theta}\mathbb{E}^{\mathbb{Q}}[X_1+X_2]$$
$$=\sup\Big\{v_{-1}(\beta_1-1)C_{-1,1}+v_0\beta_0\beta_1:(\beta_0,\sigma_0,\beta_1,\sigma_1)\in\Theta\Big\}$$
$$=:\overline{V}_0. \tag{39}$$

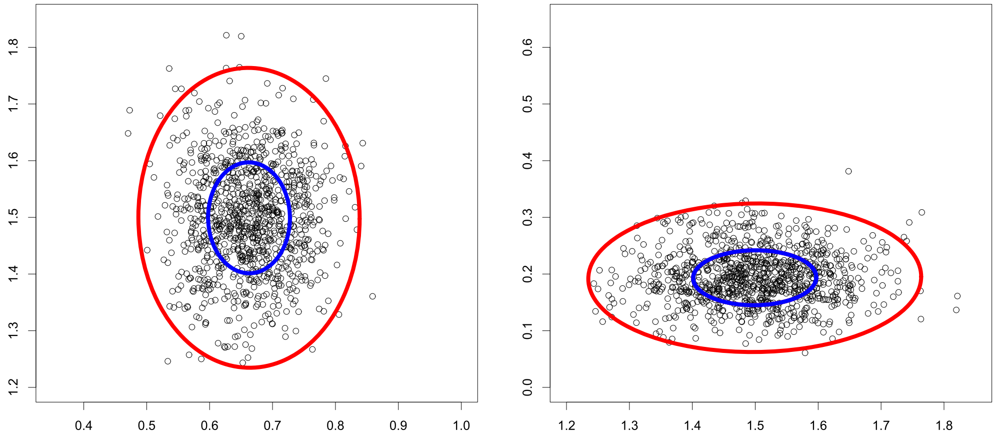

**Fig. 1.** Scatter plots of 1000 iid estimates of $(\beta_0^{\mathbb{P}}, \beta_1^{\mathbb{P}})$ (left) and of $(\beta_1^{\mathbb{P}}, \sigma_1^{\mathbb{P}})$ (right), together with boundaries of the parameter regions $\Theta_{\beta_0,\beta_1}$ (left) and $\Theta_{\beta_1,\sigma_1}$ (right) for $p = 0.1$ (blue) and for $p = 0.9$ (red).

From (10) we have the lower bound

$$V_0 \geq \sup_{\mathbb{Q}\in\mathcal{Q}_\Theta} \inf_{\tau\in\mathcal{S}_{1,T+1}} \mathbb{E}_0^{\mathbb{Q}}\left[\sum_{s=1}^{\tau-1} X_s + R_{\tau-1}\right] =: \underline{V}_0.$$

In the setting of Section 3, for each $\theta \in \Theta$, with $V_T^\theta = R_T^\theta = 0$, we solve the backward recursion

$$R_t^\theta = \rho_t(-X_{t+1} - V_{t+1}^\theta),$$
$$V_t^\theta = R_t^\theta - \mathbb{E}_t^{\mathbb{Q}_\theta}[(R_t^\theta - X_{t+1} - V_{t+1}^\theta)^+],$$

and then compute

$$\underline{V}_0 = \sup_{\theta\in\Theta} V_0^\theta.$$

Notice that $V_t^\theta, R_t^\theta$ corresponds to the quantities $V_t, R_t$ in the special case $\mathcal{Q} = \{\mathbb{Q}_\theta\}$. Computing $\underline{V}_0$ is simpler than computing $V_0$ since the former involves just one optimization over the parameter set $\Theta$ rather than $T$ nested optimizations for the latter.

We now demonstrate how $\underline{V}_0$ is computed in the current Gaussian setting. As shown above $R_1^\theta = v_0(\beta_1^{\mathbb{P}} - 1)C_{0,1} + \sqrt{v_0}\sigma_1^{\mathbb{P}} c$ (which does not depend on $\theta$) and

$$C_1^\theta = \mathbb{E}_1^{\mathbb{Q}_\theta}\left[\left(\rho_1(-X_2) - X_2\right)^+\right] = \mathbb{E}_1^{\mathbb{Q}_\theta}[\left(a(\theta, C_{0,1}) - b(\theta)e_2^\theta\right)^+],$$

where $e_2^\theta \sim N(0,1)$ with respect to $\mathbb{Q}_\theta$ and independent of $\mathcal{F}_1$, and

$$\begin{aligned} a(\theta, C_{0,1}) &= -\mathbb{E}_1^{\mathbb{Q}_\theta}[X_2] + \mathbb{E}_1^{\mathbb{P}}[X_2] + \mathrm{Var}_1^{\mathbb{P}}(X_2)^{1/2}\rho_1(e_2^{\mathbb{P}}) \\ &= v_0(\beta_1^{\mathbb{P}} - \beta_1)C_{0,1} + \sqrt{v_0}\sigma_1^{\mathbb{P}} c, \\ b(\theta) &= \mathrm{Var}_1^{\mathbb{Q}_\theta}(X_2)^{1/2} = \sqrt{v_0}\sigma_1. \end{aligned}$$

Straightforward calculations show that

$$\begin{aligned} \mathbb{E}_1^{\mathbb{Q}_\theta}[\left(a(\theta, C_{0,1}) - b(\theta)e_2^\theta\right)^+] &= a(\theta, C_{0,1})\Phi\left(\frac{a(\theta, C_{0,1})}{b(\theta)}\right) \\ &\quad + b(\theta)\varphi\left(\frac{a(\theta, C_{0,1})}{b(\theta)}\right) \\ &=: g(\theta, C_{0,1}). \end{aligned}$$

Consequently,

$$\begin{aligned} X_1 + V_1^\theta &= X_1 + R_1^\theta - C_1^\theta \\ &= v_{-1}(\beta_1 - 1)C_{-1,1} + \sqrt{v_{-1}}\sigma_1\varepsilon_{-1,2}^\theta + \sqrt{v_0}\sigma_1^{\mathbb{P}} c \\ &\quad + \beta_1^{\mathbb{P}}\left(v_0\beta_0 + \sqrt{v_0}\sigma_1\varepsilon_{0,1}^\theta\right) \\ &\quad - g\left(\theta, v_0\beta_0 + \sqrt{v_0}\sigma_1\varepsilon_{0,1}^\theta\right) \\ &= v_{-1}(\beta_1^{\mathbb{P}} - 1)C_{-1,1} + \sqrt{v_{-1}}\sigma_1^{\mathbb{P}}\varepsilon_{-1,2} + \sqrt{v_0}\sigma_1^{\mathbb{P}} c \\ &\quad + \beta_1^{\mathbb{P}}\left(v_0\beta_0^{\mathbb{P}} + \sqrt{v_0}\sigma_1^{\mathbb{P}}\varepsilon_{0,1}\right) \\ &\quad - g\left(\theta, v_0\beta_0^{\mathbb{P}} + \sqrt{v_0}\sigma_1^{\mathbb{P}}\varepsilon_{0,1}\right) \end{aligned}$$

from which $R_0^\theta = \rho_0(-X_1 - V_1^\theta)$ can be estimated with arbitrary accuracy by simulating iid copies of $X_1 + V_1^\theta$ with respect to $\mathbb{P}$ and computing the empirical estimate, and $C_0^\theta = \mathbb{E}^{\mathbb{Q}_\theta}[(R_0^\theta - X_1 - V_1^\theta)^+]$ can be estimated similarly by simulating iid copies with respect to $\mathbb{Q}_\theta$ and approximating the expectation by the empirical mean. Finally,

$$\underline{V}_0 = \sup_{\theta\in\Theta}(R_0^\theta - C_0^\theta) = \sup_{\theta\in\partial\Theta}(R_0^\theta - C_0^\theta).$$

Table 1 shows numerical values for lower bounds $\underline{V}_0$ and for upper bounds $\overline{V}_0$. These values are based on $v_{-1} = v_0 = 1$, $C_{-1,1} = \beta_0^{\mathbb{P}}$, $\rho_t = \mathrm{V@R}_{t,q}$ with $q = 0.005, 0.01, 0.05, 0.10$ and parameters sets $\Theta$ of varying size corresponding to $r^2 \leq F_{\chi^2(4)}^{-1}(p)$ with $p = 0.1, 0.5, 0.9$. The main message of Table 1 is that the intervals $(\underline{V}_0, \overline{V}_0)$ are very narrow for $q$ small and therefore the upper bound $\overline{V}_0$ is an accurate estimate of $V_0$ when $q$ is small. Notice that the upper bound is both easily computed and has attractive theoretical properties.

### 5.2. Case 2: computing $V_0$ and an upper bound for $V_0$

In this case $\mathcal{Q} = \widetilde{\mathcal{Q}}_\Theta$ and the general lower bound $\underline{V}_0$ coincides with $V_0$ and therefore its computation by backward recursion is somewhat involved. However, the upper bound is still fairly straightforward to compute. Notice that the lower bound computed for Case 1 is a lower bound for $V_0$ in the current Case 2 since $\mathcal{Q}_\Theta \subset \widetilde{\mathcal{Q}}_\Theta$.

We begin by computing the upper bound using the law of iterated expectations, extended to the multiple prior setting, and Theorem 6:

$$\overline{V}_0 = \sup_{\mathbb{Q}\in\widetilde{\mathcal{Q}}_\Theta} \mathbb{E}_0^{\mathbb{Q}}[X_1 + X_2]$$

$$= \sup_{\mathbb{Q}\in\widetilde{\mathcal{Q}}_\Theta} \mathbb{E}_0^{\mathbb{Q}}[X_1 + \operatorname{ess\,sup}_{\mathbb{Q}'\in\widetilde{\mathcal{Q}}_\Theta} \mathbb{E}_1^{\mathbb{Q}'}[X_2]]$$
$$= \sup_{\mathbb{Q}\in\mathcal{Q}_\Theta} \mathbb{E}_0^{\mathbb{Q}}[X_1 + \operatorname{ess\,sup}_{\mathbb{Q}'\in\mathcal{Q}_\Theta} \mathbb{E}_1^{\mathbb{Q}'}[X_2]].$$

Notice that, with $\beta_{1,\min} > 1$,

$$\operatorname{ess\,sup}_{\mathbb{Q}'\in\mathcal{Q}_\Theta} \mathbb{E}_1^{\mathbb{Q}'}[X_2] = v_0(\beta_{1,\max}-1)C_{0,1}\mathbb{I}_{\{C_{0,1}\ge 0\}} + v_0(\beta_{1,\min}-1)C_{0,1}\mathbb{I}_{\{C_{0,1}<0\}},$$

where $\beta_{1,\max} := \max\{\beta_1 : (\beta_0,\sigma_0,\beta_1,\sigma_1)\in\Theta\}$ and similarly for $\beta_{1,\min}$. Therefore,

$$\overline{V}_0 = \sup_{(\beta_0,\sigma_0,\beta_1,\sigma_1)\in\Theta} \Big( v_{-1}(\beta_1-1)C_{-1,1} + v_0\beta_0 + v_0(\beta_{1,\max}-1)\big(\beta_0+\sigma_0\Phi(-\beta_0/\sigma_0)\big) - v_0(\beta_{1,\min}-1)\sigma_0\Phi(-\beta_0/\sigma_0)\Big)$$

$R_1$ is calculated explicitly as above. Computing $C_1$ means computing

$$C_1 = \operatorname{ess\,sup}_{\theta_1\in\partial\Theta_{\beta_1,\sigma_1}} g(\theta_1, C_{0,1}),$$

where, with some abuse of notation, we consider $g$ to be defined for parameters $\theta_1 \in \Theta_{\beta_1,\sigma_1}$ rather than $\theta\in\Theta$. In practice, this means determining a function $h:\mathbb{R}\to\mathbb{R}$ such that $h(C_{0,1}^k) = \max_{\theta_1\in\partial\Theta_{\beta_1,\sigma_1}} g(\theta_1, C_{0,1}^k)$ for suitably many simulated iid copies $C_{0,1}^1,\dots,C_{0,1}^n$ of $C_{0,1}$ and approximating $C_1 \approx h(C_{0,1})$. Given the choice of $h$, $R_0 = \rho_0(-X_1 - R_1 + C_1)$ is approximated by its empirical estimate based on simulated iid copies with respect to $\mathbb{P}$ of

$$v_{-1}(\beta_1^{\mathbb{P}}-1)C_{-1,1} + \sqrt{v_{-1}}\sigma_1^{\mathbb{P}}\varepsilon_{-1,2} + \sqrt{v_0}\sigma_1^{\mathbb{P}}c + \beta_1^{\mathbb{P}}\left(v_0\beta_0^{\mathbb{P}} + \sqrt{v_0}\sigma_1^{\mathbb{P}}\varepsilon_{0,1}\right) - h\left(v_0\beta_0^{\mathbb{P}} + \sqrt{v_0}\sigma_1^{\mathbb{P}}\varepsilon_{0,1}\right)$$

Similarly, $C_0$ is approximated by, for each $\theta$ in a dense subset of $\partial\Theta$, simulating iid copies with respect to $\mathbb{Q}_\theta$ of

$$v_{-1}(\beta_1-1)C_{-1,1} + \sqrt{v_{-1}}\sigma_1\varepsilon_{-1,2}^{\theta} + \sqrt{v_0}\sigma_1^{\mathbb{P}}c + \beta_1^{\mathbb{P}}\left(v_0\beta_0 + \sqrt{v_0}\sigma_1\varepsilon_{0,1}^{\theta}\right) - h\left(v_0\beta_0 + \sqrt{v_0}\sigma_1\varepsilon_{0,1}^{\theta}\right),$$

estimating $\mathbb{E}^{\mathbb{Q}_\theta}[(R_0 - X_1 - R_1 + C_1)^+]$ by the empirical mean, and computing the minimum of these expectations over the $\theta$ values. Finally, $V_0$ is estimated by the difference of the estimates of $R_0$ and $C_0$.

Table 1 shows numerical values for lower bounds $\underline{V}_0$ and for upper bounds $\overline{V}_0$ with the same parameter values as those considered for Case 1. Similarly to Case 1, the intervals $(\underline{V}_0, \overline{V}_0)$ are very narrow for $q$ small and therefore the upper bound $\overline{V}_0$ is an accurate estimate of $V_0$ when $q$ is small.

**Table 1**
Case 1 and Case 2: lower and upper bounds $(\underline{V}_0, \overline{V}_0)$ rounded to three decimals, where the size of the parameter uncertainty region is determined by $r^2 \le F^{-1}_{\chi^2(4)}(p)$ and $\rho_t = \mathrm{V@R}_{t,q}$. Empirical estimates were based on iid samples of size $10^5$.

| Case 1 | $p=0.1$ | $p=0.5$ | $p=0.9$ |
|---|---|---|---|
| $q=0.10$ | (1.452, 1.491) | (1.562, 1.624) | (1.686, 1.787) |
| $q=0.05$ | (1.473, 1.491) | (1.592, 1.624) | (1.730, 1.787) |
| $q=0.01$ | (1.490, 1.491) | (1.618, 1.624) | (1.772, 1.787) |
| $q=0.005$ | (1.491, 1.491) | (1.622, 1.624) | (1.780, 1.787) |

| Case 2 | $p=0.1$ | $p=0.5$ | $p=0.9$ |
|---|---|---|---|
| $q=0.10$ | (1.470, 1.513) | (1.595, 1.666) | (1.734, 1.856) |
| $q=0.05$ | (1.491, 1.513) | (1.628, 1.666) | (1.786, 1.856) |
| $q=0.01$ | (1.509, 1.513) | (1.656, 1.666) | (1.835, 1.856) |
| $q=0.005$ | (1.511, 1.513) | (1.661, 1.666) | (1.845, 1.856) |

## Declaration of competing interest

There are no competing interests.

## Data availability

No data was used for the research described in the article.

## Acknowledgement

We thank the two anonymous reviewers for constructive criticism that led to an improvement of the paper. Filip Lindskog acknowledges financial support from the Swedish Research Council, Project 2020-05065. Julie Thøgersen was supported by a research grant (34444) from Villum Fonden.

## References

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Ku, H., 2007. Coherent multiperiod risk adjusted values and Bellman's principle. Annals of Operations Research 152, 5–22.
Artzner, P., Eisele, K.-T., Schmidt, T., 2020. No arbitrage in insurance and the QP-rule. Available at SSRN: https://ssrn.com/abstract=3607708. https://doi.org/10.2139/ssrn.3607708.
Barigou, K., Chen, Z., Dhaene, J., 2019. Fair dynamic valuation of insurance liabilities: merging actuarial judgement with market- and time-consistency. Insurance. Mathematics & Economics 88, 19–29.
Barigou, K., Dhaene, J., 2019. Fair valuation of insurance liabilities via mean variance hedging in a multi period setting. Scandinavian Actuarial Journal 2019, 163–187.
Bion-Nadal, J., 2008. Dynamic risk measures: time consistency and risk measures from BMO martingales. Finance and Stochastics 12, 219–244.
Cheridito, P., Kupper, M., Delbaen, F., 2006. Dynamic monetary risk measures for bounded discrete-time processes. Electronic Journal of Probability 11, 57–106.
Cheridito, P., Kupper, M., 2009. Recursiveness of indifference prices and translation-invariant preferences. Mathematics and Financial Economics 2, 173–188.
Cheridito, P., Kupper, M., 2011. Composition of time-consistent dynamic monetary risk measures in discrete time. International Journal of Theoretical and Applied Finance 14, 137–162.
Deelstra, G., Devloder, P., Gnameho, K., Hieber, P., 2020. Valuation of hybrid financial and actuarial products in life insurance by a novel three-step method. ASTIN Bulletin 50, 709–742.
Delong, L., Dhaene, J., Barigou, K., 2019. Fair valuation of insurance liability cash-flow streams in continuous time: theory. Insurance. Mathematics & Economics 88, 196–208.
European Commission, 2015. Commission Delegated Regulation (EU) 2015/35 of 10 October 2014. Official Journal of the European Union.
Detlefsen, K., Scandolo, G., 2005. Conditional and dynamic convex risk measures. Finance and Stochastics 9, 539–561.
Engelage, D., 2011. Optimal stopping with dynamic variational preferences. Journal of Economic Theory 146, 2042–2074.
Engsner, H., Lindensjö, K., Lindskog, F., 2020. The value of a liability cash flow in discrete time subject to capital requirements. Finance and Stochastics 24, 125–167.
Engsner, H., Lindholm, M., Lindskog, F., 2017. Insurance valuation: a computable multi-period cost-of-capital approach. Insurance. Mathematics & Economics 72, 250–264.
Föllmer, H., Schied, A., 2016. Stochastic Finance: An Introduction in Discrete Time, 4th edn. Walter de Gruyter.
Grosen, A., Jørgensen, P.L., 2002. Life insurance liabilities at market value: an analysis of insolvency risk, bonus policy, and regulatory intervention rules in a barrier option framework. The Journal of Risk and Insurance 69, 63–91.
Hieber, P., Natolski, J., Werner, R., 2019. Fair valuation of cliquet-style return guarantees in (homogeneous and) heterogeneous life insurance portfolios. Scandinavian Actuarial Journal 2019, 478–507.

Kallenberg, O., 2002. Foundations of Modern Probability, 2nd edn. Springer.
Lindholm, M., Lindskog, F., Wahl, F., 2017. Valuation of non-life liabilities from claims triangles. Risks 5 (3).
Krätschmer, V., Ladkau, M., Laeven, R.J.A., Schoenmakers, J.G.M., Stadje, M., 2018. Optimal stopping under uncertainty in drift and jump intensity. Mathematics of Operations Research 43 (4), 1177–1209.
Mack, T., 1993. Distribution-free calculation of the standard error of chain ladder reserve estimates. ASTIN Bulletin 23, 213–225.
Malamud, S., Trubowitz, E., Wüthrich, M.V., 2008. Market consistent pricing of insurance products. ASTIN Bulletin 38, 483–526.
Möhr, C., 2011. Market-consistent valuation of insurance liabilities by cost of capital. ASTIN Bulletin 41, 315–341.
Pelsser, A., Stadje, M., 2014. Time-consistent and market-consistent evaluations. Mathematical Finance 24, 25–65.
Riedel, F., 2004. Dynamic coherent risk measures. Stochastic Processes and Their Applications 112, 185–200.
Riedel, F., 2009. Optimal stopping with multiple priors. Econometrica 77, 857–908.
Rockafellar, R.T., Uryasev, S., Zabarankin, M., 2006. Generalized deviations in risk analysis. Finance and Stochastics 10, 51–74.
Rosazza Gianin, E., 2006. Risk measures via g-expectations. Insurance. Mathematics & Economics 39, 19–34.
Tsanakas, A., Wüthrich, M.V., Černý, A., 2013. Market value margin via mean-variance hedging. ASTIN Bulletin 43, 301–322.
Wüthrich, M.V., Embrechts, P., Tsanakas, A., 2011. Risk margin for a non-life insurance runoff. Statistics and Risk Modeling 28, 299–317.
Wüthrich, M.V., Merz, M., 2013. Financial Modeling, Actuarial Valuation and Solvency in Insurance. Springer.