CVaR based pricing and hedging in Unit-Linked insurance products

Andreas Würth

March 11, 2008

Abstract

This paper treats the question how to price and hedge unit-linked insurance products such that the claim becomes acceptable using CVaR as risk measure. We connect our question to the results on expected shortfall by Föllmer/Leukert and prove some extensions of those results, in order to make the theorem also applicable to discrete insurance outcomes as is typical for life insurance. We propose a way to obtain upper bounds for the minimal price required by discretization of the time when information about the insurance process arrives, and prove a convergence result. Finally, we show how to apply those results for obtaining price and hedge in the case of a unit-linked survival insurance, where we obtain analytical formulas for the shape of the optimal hedge.

1 Introduction

Unit linked insurance products become more and more popular in insurance industry, because they combine the classical coverage against risks such as death, longlife and disability in life insurance with the possible chance of large capital earnings that traditionally banks offer. For the insurer, such a product has the advantage that he has no exposure to the interest rate risk.

The pricing of such insurance products needs a combination of classical actuarial principles as well as principles from financial mathematics. Such combinations have been treated in [1], where the focus was mainly on pricing using a standard deviation principle, but also some hints for a general utility function. In general, financial valuation principles base on a replication of a claim, whereas in insurance, a risk-loading is charged, because the claims cannot be hedged. For unit-linked insurance products, one can assume that a full hedge is not possible, because there is a nonhedgeable component. However, because of the financial component of those products, at least a partial hedge should be possible. This partial hedge should be done in a way that the risk, however defined, is minimized using this strategy. The aim of this paper is to show how such a risk-minimizing strategy looks like and how it may be found.

In life insurance industry, most popular risk measures are value at risk and conditional value at risk or expected shortfall. We restrict therefore in this paper to those measures, in particular essentially to expected shortfall. It has been shown in [3] and [4] that minimization of expected shortfall is connected

to the Neyman-Pearson theory [5], and for some specific cases they develope analytic formulas for the case where the market is complete, and some general principles for the hedging of a volatility jump at a specific time. We further develope those results for the specific case of unit-linked insurance products.
Risk-minimizing strategies for unit-linked insurance products have already been treated in [2], where risk-minimizing is understood in the sense of Föllmer and Schweizer. Other papers of the same author consider the variance as definition of risk. Minimization of value at risk in unit-linked insurance products has been looked at by [7], where the authors have mainly focussed on the case with only one insured person and information about the insurance process at the end of the insurance period, and in this way obtained analytic formulas. Including the diversification effect among insured persons, they obtain only bounds.
In our paper, we consider expected shortfall as risk measure. In contrast to [7], we are interested in a risk minimizing strategy when the diversification effect is included, as well as information about the insurance process may be revealed before the terminal payoff time. Furthermore, we are interested not only in analytical formulas, but also in questions of numerical tractability.
Some general papers about coherent risk measures and the dynamic programming principle are [8] and [9], where the authors mainly use the duality result stated in [10], that is that the risk measure can be written as a supremum over a set of test measures. Assuming that those test measures are equivalent and the Girsanov processes due to this density processes are known, one can take them as control processes. However, this is only a way to obtain the risk measure; the minimization problem would require a further minimization, which is a more delicate numerical issue. Moreover, this technique is not appropriate to measures like conditional value at risk, because the test measures corresponding to it are mostly not equivalent. Furthermore, it is not clear how to obtain the set of density processes corresponding to this risk measure numerically, even if theoretically it is clear what is meant.
We went therefore another approach, following [12], which gave us again a dynamic programming problem, which can in principle be solved by partial differential equations technique, as proposed for example in [13]. However, the terminal value function for this problem is neither differentiable (actually it has a kink) nor strictly concave. Furthermore, one has to take not only the assets themselfs and the insurance variables, but also the actual wealth as state variable, because in contrast to exponential hedging, the value function depends also on the current wealth. This inceases the amount of dimensions even for simple problems, and furthermore, leads to the fact that the conditions for convergence stated in [13] are not satisfied. Even if we obtain results, we should also be able to verify our numerical approximation.
We will propose another approach for obtaining lower bounds of the value function, which follows again the results of [4]. Indeed, we perform a discretization not in time itself, but in information time about the insurance process. We will show that this discretization gives always lower bounds for the value function, and those bounds converge to the optimal value function if time step between two points of information decrease to zero. Even if we still need a grid in wealth and all state variables, it should be possible to apply this method also using a

nonlinear grid or by simulation, which makes it more likely that this method also works for higher dimensions.
The structure of the paper is as follows: In section 2, we formulate the model, as well as the optimization problem. In section 3, we treat the situation of insurance information only at the end of the period, and we proof a generalization of the results of [4], suitable for the situation in insurance. This result allows us to know where to find the optimal payoff, and to obtain analytical formulas for some cases, or to solve the problem numerically by a Fast Legendre Transformation method. A specific example of a unit linked survival insurance is given in section 4, where anaytical formulas are obtained. We will construct also the worst-case martingale measure in the sense of Neyman-Pearson [5], also stated in [3]. In section 5, we will extend to the situation where information about the insurance process is available before the terminal time, and we will prove the continuous-time information limit. Section 6 then treats more detailed the numerical issues, again for the example of a unit-linked survival insurance. Section 7 concludes.

2 Problem specification and general statements

2.1 Insurance model and problem specification

Our insurance model consists of a vector-valued stochastic process Z_t representing the insurance state process. This process is assumed to be Markovian. The states of the financial market are represented by another vector-valued process X_t. It is assumed that this process is generated by a multivariate Brownian motion, and that this market is complete, in the sense that every contingent claim $F(X_T)$ can be replicated by a suitable trading strategy

$$F(X_T) = S_0 + \int_0^T \pi_t dS_t$$

where S_t is the vector-valued process representing the available financial assets. It is assumed that this process is a vector-valued function of time and the financial state variables, that is

$$S_t = H(t, X_t)$$

with H a measurable function. Apart from those assets S_t, it is assumed that there do not exist further assets in the market.
It is assumed furthermore that $\sigma(X)$, the generating sigma-algebra of the financial market process, is independent of $\sigma(Z)$, the insurance process.
The option payoff due to a unit-linked insurance product at the terminal time T is given by a nonnegative product-measurable function $g(X_T, Z_T)$, depending on the financial market as well as on the insurance process.

Remarks:

- By extending the state space, it is always possible that the payoff depends on some states at $t < T$. Therefore, to restrict to payoffs depending only on states at time T is not really a restriction.
- The restiction which has been made is that payoffs can only be made at time T, even if they may depend on earlier times. This has to be made because we aim to minimize the expected shortfall at the terminal time T. It is in general not clear how to define the expected shortfall if there are different payoff times. In some specific examples, it may be sensible to divide the payments at all times by a numéraire (a reasonable choice may be a zero bond with expiry at time T), and take the expected shortfall at the fixed time T. This situation is also covered by our model.

The CVaR pricing rule is now the following: Find the minimal capital V_0 such that there exists a self-financing predictable strategy π with

$$CVaR\left[(g(X_T, Z_T) - Y_T^{\pi})\right] \leq 0 \tag{1}$$

and such that the wealth process Y_t^{π}, defined as

$$Y_t = V_0 + \int_0^T \pi dS$$

remains nonnegative for all t. Furthermore, find the corresponding strategy π.

Remarks:

- Due to the general theory about coherent risk measures [10], equation (1) means that the risk is acceptable for the insurance company.
- The nonnegativity is no restriction with respect to the more general case that the capital remains bounded from below by a value $-cB_t$, where c is a constant and B is a numéraire. Indeed, adding capital cB_0 to the initial available capital, and cB_T to the terminal payoff, the restriction of a nonnegative wealth process Y_t of this modified problem is the same as the restriction that $Y_t \geq cB_t$ of the original problem.
- Conversely, for a fixed initial capital V_0, one can also ask the question what is the minimal possible CVaR and the corresponding hedging strategy. This question is sensible if the market is competitive and an insurance company is not able to price independently its portfolio.

2.2 Market consistent CVaR

The pricing rule from (1) does not necessarily give a market consistent price. Indeed, in a Black-Scholes model, if we have the replicable claim $g(X_T) = 1_{X_T<c}$ with c a very small constant, we may obtain a CVaR smaller 0 at the risk-neutral price of this option. Similarly, if we define the pure risk premium as

the expectation of the claim under the measure Q^*, which gives the original probabilities for the insurance process and the risk-neutral ones for the financial process, one may obtain a negative risk premium.
This is a problem of the CVaR risk measure. In practice, a negative risk loading happens very seldom, and it will never happen in our examples that we calculate. To be strict and to guarantee market consistency, one has to change slightly the risk measure. We propose to do this by adding the measure Q^* to the set of test measures in the dual representation of CVaR. This gives a coherent risk measure, and the new pricing rule is market consistent and is nothing else than what practitionars applying the CVaR criterion for pricing would do most likely: When the price obtained by (1) would have a negative risk loading, they would set the risk loading to 0, otherwise, they would let the price as obtained by the CVaR criterion.
We denote this new risk measure as market consistent CVaR. In practice, in particular in all of our examples, there is no difference between this and the CVaR.

2.3 Connection to minimization of Expected Shortfall

It follows from [12] that the conditional value at risk of a random variable X at a level β us given by

$$CVar(X) = \min_{\alpha} \left(\alpha + \frac{1}{1-\beta} E\left[(X-\alpha)^+\right] \right) \tag{2}$$

Using this, the problem from last section can be reformulated to the following: Find the minimal initial capital V_0 such that there exist a allowed strategy π and a parameter α with

$$f(\alpha, \pi; V_0) \leq -\alpha(1-\beta) \tag{3}$$

where $f(\alpha, \pi; V_0)$ is given by

$$f(\alpha, \pi; V_0) := E\left[\left((g(X_T, Z_T) - \alpha)^+ - Y_T^\pi \right)^+ \right] \tag{4}$$

Remarks:

- From (3) and (4), it is clear that α must be nonpositive. It follows that $(g(X_T, Z_T) - \alpha) \geq 0$ is always satisfied if the insurance claim is nonnegative.

- We have that $((g(X_T, Z_T) - \alpha)^+ - Y_T^\pi)^+ = (g(X_T, Z_T) - \alpha - Y_T^\pi)^+$, because of the nonnegativity of Y_T^π. The reason why we write the expectation as in (4) will become clear later, when we will connect this problem to the statements in [3] and [4].

- By [12], the minimum in α is always attained. It follows that the minimum in V_0 under condition (1) is attained if the minimum $V_0(\alpha)$ under condition (3) is attained for all fixed α.

For fixed α, we can define $V_0(\alpha)$ as the minimal initial capital such that there exists a strategy π which satisfies (3).

Proposition 2.1. *The following is true:*

1. *For each α, the minimum $V_0(\alpha)$ is attained.*
2. *The function $V_0(\alpha)$ is convex in α.*
3. *If α^* minimizes $V_0(\alpha)$ and π^* is the strategy which minimizes $V_0(\alpha^*)$ for the given α^*, then π^* is the strategy which makes the claim at initial capital $V_0(\alpha^*)$ acceptable due to criterion (1).*

Proof. For the first statement, we follow the arguments of [3]. For each α, we can define a measurable random variable $\phi \in [0,1]$ with

$$\left((g(X_T, Z_T) - \alpha)^+ - Y_T^\pi\right)^+ = (1-\phi)(g(X_T, Z_T) - \alpha)^+$$

which has been denoted the success ratio in [3]. The problem becomes now the one to minimize

$$sup_{Q \in \mathcal{Q}} E^Q[\phi(g(X_T, Z_T) - \alpha)^+]$$

where $\mathcal{Q}$ is the set of all equivalent martingale measures, under the condition that

$$\hat{E}[\phi] \geq 1 + \alpha(1-\beta)$$

where $\hat{E}$ denotes the expectation under the measure $\hat{P}$, defined by

$$\frac{d\hat{P}}{dP} = \frac{(g(X_T, Z_T) - \alpha)^+}{E[(g(X_T, Z_T) - \alpha)^+]}$$

Existence of an optimal ϕ follows now by the same argument as in [3], and also the optimizing strategy which relies on the optional decomposition theorem. For the second statement, let V_1 be the mimimal required capital for α_1 and V_2 the same for α_2, where α_1 and α_2 are arbitrary real numbers, and π_1 and π_2 the corresponding strategies. Then (3) is satisfied for (α_1, π_1, V_1) as well as for (α_2, π_3, V_2), and for an arbitrary $t \in [0,1]$ we have

$$tf(\alpha_1, \pi_1; V_1) + (1-t)f(\alpha_2, \pi_2; V_2) \leq -(t\alpha_1 + (1-t)\alpha_2)(1-\beta)$$

By the convexity of the function $x \to x^+$, it follows that the left-hand side is larger or equal than $f(t\alpha_1 + (1-t)\alpha_2, t\pi_1 + (1-t)\pi_2; tV_1 + (1-t)V_2)$, so that

$$f(t\alpha_1 + (1-t)\alpha_2, t\pi_1 + (1-t)\pi_2; tV_1 + (1-t)V_2) \leq -(t\alpha_1 + (1-t)\alpha_2)(1-\beta)$$

It follows that at capital $tV_1 + (1-t)V_2$, there exists a strategy such that this equation is satisfied, and the minimal capital must therefore be smaller or equal. The third statement is obvious. $\square$

It follows that we can firstly minimize the required capital $V_0(\alpha)$ for a fixed α, and after minimize this expression with respect to α. But the first one is a problem of the type discussed in [4], namely the minimization of the capital required provided an expected shortfall constraint. We can therefore apply the considerations made in this paper.

It will sometimes be easier to consider the related problem, namely to minimize $f(\alpha, \pi; V_0)$ at a given initial capital V_0. It is clear that

$$f_{min}(\alpha, V_0) := \min_{\pi} f(\alpha, \pi; V_0) \tag{5}$$

is a nonincreasing function in V_0 for given α. If we have $f_{min}(\alpha, V_0)$ for all α and V_0, we can take the minimal V_0 such that

$$f_{min}(\alpha, V_0) \leq -\alpha(1-\beta)$$

An advantageous situation occurs if f_{min} is continuous in V_0. In this case, we can replace the inequality sign by equality.
If we aim to minimize the CVaR at a given initial capital V_0, we can again apply equation (5). By [12], the function

$$\alpha + \frac{1}{1-\beta} f_{min}(\alpha, V_0)$$

is convex in α, and we can again make the minimization over all values of α.

It becomes clear that the essential problem is (5), which is a problem of minimizing expected shortfall in the sense of [4], and from which everything else follows. In the sequel, we will therefore focus on this problem. As in [3] and [4], we reformulate the problem of minimizing expected shortfall as a problem of maximizing a state-dependent utility function, that is we write

$$E[(g(X_T, Z_T) - Y_T^{\pi})^+] = E[g(X_T, Z_T)] - E[g(X_T, Z_T) \wedge Y_Y^{\pi}]$$

We can therefore, instead of minimizing equation (5), maximize

$$E[(g(X_T, Z_T) - \alpha)^+ \wedge Y_T^{\pi}] \tag{6}$$

under the condition that

$$B_0 E^Q[Y_T^{\pi}] \leq V_0$$

for all equivalent martingale measures Q, where B_0 is the value of the zero bond with expiry time T, which is taken here as numéraire.
In the sequel, we will take expression (6) as objective function.

2.4 Insurance information at the end of the period

A special case occurs if all information about the insurance states is only available at the end of the period, that is at time T. In this case, we can refer to similar considerations as in [4]. We will consider this special case firstly, and

generalize it later in the paper.
For any strategy we followed up to time T^-, the objective function at time T^- is given by

$$E[g(X_T, Z_T) \wedge Y_T^\pi | X_{T^-}, Y_{T^-}]$$

and by the Markov property and the predictability of X_t and Z_t, this expression is equal to $h(X_T, Y_T^\pi)$, where

$$h(x, y) := \int (g(x, z) \wedge y) \, dP(z) \tag{7}$$

where $dP(z)$ is the distribution function of Z, which is by assumption independent of $\mathcal{F}_{T^-}$, and therefore in particular independent of the strategy. The function h is concave in the second argument, and the optimization problem is now

$$\max_\pi E[h(X_T, Y_T^\pi)] \tag{8}$$

subject to

$$B_0 E^Q[h(X_T, Y_T^\pi)] \leq V_0$$

Because this problem is independent of $\sigma(Z)$, we are in a complete model, and the martingale measure Q is unique. In life insurance, we can mostly assume that

$$u(x) := \sup_z g(x, z) < \infty$$

and that

$$V_{sup} := B_0 E^Q[u(X_T)] < \infty$$

In this case, a superhedge is possible at a finite initial capital, and we can apply to a large extent the theory developed in [3] or [4]. However, some adjustments and extensions are necessary:

1. Our value function $h(x, y)$ is concave but not strictly concave, nor differentiable. We have to generalize the theory for covering also this situation.

2. It has already been indicated in [3] how to proceed if a new realization of the nonhedgeable claim takes place before time T. We will follow essentially those arguments. However, we are also interested in the limiting case where the new information about Z_t arrives continuously.

3. In [4], a risk-free asset with interest rate $r = 0$ has been assumed to exist. We do not want to assume this, and therefore we have to discount by a numéraire, which will be in most cases the zero bond up to time T. We want to see whether or not this causes additional problems.

For problem 1, we will prove a theorem which is an extension of the considerations of [4], and which gives the required generalization. With this theorem, we will be able to know where to find the optimal solution in the situation where information is only available at the end of the period. For this specific case, problem 3 is trivial, because B_0 is then only a number. For the extension to the general case, we will then treat problems 2 and 3.

3 Example: Unit-linked survival insurance

The idea of this section is to show the way how to calculate the price and hedge determined by the CVaR criterion using a simple example.
We are treating here the case of a unit-linked survival insurance, where the stock price S_T is paid at time T if the person is still alive, and nothing if the person has died before this time. The process S_t is assumed to follow a geometric Brownian motion, that is

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{9}$$

and the amount of survivors a binomial distribution. It is assumed that the insurance outcome is independent of the Brownian motion process. For simplicity, we assume the risk-free interest rate to be 0. If n is the amount of insured persons, and p the probability of surving, then the amount of survivors N_T is

$$N_T \sim BIN(n,p) \tag{10}$$

The total insurance payoff is

$$g(S_T, N_T) = S_T N_T \tag{11}$$

It is furthermore assumed that the information about the survivors is firstly reveald at time T. We may therefore apply equation (7) for obtaining the expected value at time T^- for a given final stock price and capital. For large values of n, it seems reasonable to approximate the amount of survivors by a normal distribution, that is

$$N_T \sim N(np, np(1-p)) \tag{12}$$

Integrating out this normal random variable due to (7) it follows, for a given α and initial capital V_0,

$$\begin{aligned} \alpha(x,y) &= (npx-\alpha)\left(\Phi\left(\frac{\frac{y+\alpha}{x}-np}{np(1-p)}\right) - \Phi\left(\frac{\frac{\alpha}{x}-np}{np(1-p)}\right)\right) \\ &\quad + y\left(1-\Phi\left(\frac{\frac{y+\alpha}{x}-np}{np(1-p)}\right)\right) + \frac{np(1-p)x}{2\pi}\left(e^{-\frac{\left(\frac{\alpha}{x}-np\right)^2}{2np(1-p)}} - e^{-\frac{\left(\frac{y+\alpha}{x}-np\right)^2}{2np(1-p)}}\right) \end{aligned} \tag{13}$$

One can easily see that the derivative of α with respect to y is

$$\frac{\partial}{\partial y}\alpha(x,y) = 1 - \Phi\left(\frac{\frac{y+\alpha}{x} - np}{np(1-p)}\right)$$

from which it follows that α is strictly increasing, strictly concave and differentiable. For determining the optimal payoff to be hedged, we would like to apply Proposition 8.2 in [4]. However, we cannot directly do this, because we have no superhedge in the normal approximation. But this is only an approximation, and in reality, to hedge nS_T gives a superhedge. Theorem 7.1 in [4] states then that the optimal claim to be hedged is

$$v(x)_\alpha = 0 \vee \left(-\alpha + npx + np(1-p)x\Phi^{-1}\left(1-\gamma x^{-\frac{mu}{\sigma^2}}\right)\right) \wedge nx \tag{14}$$

Table 1: Minimal capital for different parameters

n	p	γ	α	V_0	*Load*
1000	0.5	0.052	-6.73	532.61	6.5%
1000	0.1	0.052	-4.04	119.56	19.6%
50	0.5	0.052	-1.50	32.29	29.2%
50	0.1	0.052	-0.90	9.37	87.4%

where we have inserted here equation (13) and applied that in model (9), the density at time T is proportional to $S_T^{-\mu}\sigma^2$. The parameter γ is determined by the budget constraint $V_0 = E^Q[v_\alpha(S_T)]$.
We want now to determine the minimal capital V_0 such that the $CVaR(N_T S_T - V_T)$ is less or equal 0 and the wealth is always nonnegative. For that, as in the previous section, we firstly determine the minimal capital $V_0(\alpha)$ for a fixed α and minimize then over all α. For a given $V_0(\alpha)$, the minimal expected shortfall must be, by the previous considerations, of the form

$$ES(\alpha) = E[(N_T S_T - \alpha)^+] - E[\alpha(S_T, v_\alpha(S_T))] \tag{15}$$

with v_α from (14) and the parameter γ determined by $V_0(\alpha) = E^Q[v_\alpha(S_T)]$. It follows that $ES(\alpha)$ is a continuous and decreasing function in γ, whereas $V_0(\alpha)$ is a continuous and increasing function in γ. The minimum $V_0(\alpha)$ is attained at the γ at which

$$ES(\alpha) = -\alpha(1-\beta) \tag{16}$$

This gives a one-to-one relationship between α and γ. Instead of minimizing $V_0(\alpha)$, we can do the minimization with respect to γ, calculating the corresponding α through (16). We did this numerically using the parameters $\mu = 0.07$, $\sigma = 0.2$, $T = 1$, $S_0 = 1$, $\beta = 0.95$. The results are in the following Table 3. The last column, *Load*, is the risk loading as percentage of the pure premium. The optimal hedge is then the complete delta hedge of the payoff $V_\alpha(S_T)$, with α and the corresponding γ given in table 3.

In typical situations of life insurance, we have however a discoutinuous distribution function for the amount of survivors. Applying (7) leads then to a piecewise linear function $\alpha(x, y)$, which is therefore not differentiable nor strictly concave. We can therefore not apply theorem 7.1. In the following section, we will prove therefore a corresponding theorem for the case where no differentiability nor strict concavity of the state-dependent utility function is needed.
Another assumption we made here was the one that information about the insurance process arrives only at the end of the period. This seems not very realistic. However, it gives an upper bound for the minimal price V_0 as well as the corresponding hedge such that the CVaR remains indeed less or equal 0. From this point of view, one can see this already as a reasonable pricing rule. For exploiting more information about the insurance process and obtaining better bounds for V_0, one can follow an idea already pointed out in [3] and consider the

situation where already at a specific time point $t_0 < T$ information arrives, and repeating this procedure for many information time points. We will work out this procedure, and also show that the value function when information arrives continuously is a limit of the situation where information arrives at finite many points, as the mesh of the partition of information points tends to 0.

4 The basic property

In this section, we prove the basic property of the optimal solution for the case when insurance information is available only at the end of the time period. This result shows then where to find the optima. It is essentially an extension of the results in [3] for functions which are not strictly concave nor differentiable.
Firstly, we state here a condition which will often be used throughout this section. We denote this condition FHFC (full hedge with finite capital), because, economically speaking, this condition means that one can make a full hedge using only a finite amount of capital. In life insurance, this condition is mostly satisfied, because the assumption that everyone survives gives the worst case.

Definition: A function $\alpha(x, y)$, is said to satisfy the FHFC condition with respect to the probability measure μ on the Borel set $\mathcal{B}(A)$, $A \subset \mathbb{R}^n$ if there exists a measurable and $\mu-$ integrable function $h(x) > 0$ such that $\sup_y \alpha(x, y) = \alpha(x, h(x))$.

Theorem 4.1. *Let $A \subset \mathbb{R}^n$ an interval, and let ν and μ be two finite equivalent measures on $\mathcal{B}(A)$.*
Let $\alpha : D := A \times [0, \infty) \to \mathbb{R}$ be a function which is concave, nondecreasing in the second argument, and satisfies the FHFC condition with respect to μ, and let $\alpha(x, h(x))$ be $\nu-$ integrable. Define $\alpha(x, y) := -\infty$ for all $y < 0$, then the concavity holds for all real numbers.
Let $v : A \to [0, \infty)$ be a function in C, where C is the set of all Borel-measurable functions

$$f : A \to \mathbb{R}$$

with the property that

$$||f|| := \sup_{x \in A} \left| \frac{f(x)}{h(x) + 1} \right| < \infty \tag{17}$$

with $h(x)$ the function from the FHFC condition.
Then the following statements are equivalent:

1. *There exists a function $\beta : D \to \mathbb{R}$ such that for each $(x, y) \in D$ $\beta(x, y)$ is a point in the superdifferential of α with respect to the second argument, and such that*

$$\beta(x, v(x)) = \gamma \frac{d\mu}{d\nu}(x) \tag{18}$$

for a constant $\gamma > 0$.

2. *The function $v(x)$ optimizes*

$$\int_A \alpha(x, f(x))d\nu(x)$$

over all functions $f \in C$ with

$$\int_A f(x)d\mu(x) \leq \int_A v(x)d\mu(x) \tag{19}$$

Remark: Economically speaking, this is a functional analytical version of the marginal utility statement, where μ is the pricing functional of f, and $\nu(\alpha(., f))$ is its utility.

Proof. We first prove the easy direction from (1) to (2). Let $f(x)$ be any nonnegative function in C. We define the function

$$f_\lambda(x) := (1-\lambda)f(x) + \lambda v(x)$$

It is clear that for all $\lambda \geq 0$, $f_\lambda \in C$ and nonnegative. By the concavity of α in the second argument, we have, for any choice $\beta(x, y)$ for the superdifferential, that

$$\alpha(x, f_\lambda(x)) - \alpha(x, f_0(x)) \leq \beta(x, f_0(x))(f(x) - v(x))\lambda$$

and therefore, because satisfies FHFC and $\alpha(x, h(x))$ is integrable, we have that $\alpha(x, f_\lambda(x)) \leq \alpha(x, h(x))$ is integrable for all $\lambda \geq 0$ and

$$\int_A \alpha(x, f(x))d\nu(x) - \int_A \alpha(x, v(x))d\nu(x) \leq \int_A \beta(x, v(x))(f(x) - v(x))d\nu(x) \tag{20}$$

Now let the superdifferential satisfy property (18). Then it follows for the right-hand side of equation (20) that

$$\int_A \beta(x, v(x))(f(x) - v(x))d\nu(x) = \gamma \left(\int_A f(x)d\mu(x) - \int_A v(x)d\mu(x) \right) \leq 0$$

where the last inequality follows if f satisfies property (19), and the integrability is again guaranteed by FHFC with respect to μ. It follows from (20) that

$$\int_A \alpha(x, f(x))d\nu(x) \leq \int_A \alpha(x, v(x))d\nu(x)$$

and therefore $v(x)$ is optimal for all nonnegative functions $f \in C$. For $f(x) < 0$ on a set with $\nu-$ positive measure, $\alpha(x, f(x)) = -\infty$ on a set with $\nu-$positive measure, and the integral is $-\infty$, which cannot be optimal.

Now let us turn in the other direction. Here we need functional analytical arguments from inifinite dimensional convex analysis. We have that C, with the norm from (17) is a convex normed vector space. Furthermore, the function F on C defined by

$$F(f) := \int_A \alpha(x, f(x))d\nu(x)$$

is a concave function, which follows easily by the concavity of α. The function $f \to \int_A f(x)d\mu(x)$ is a continuous linear functional on C. Furthermore, if $v(x)$ is not identically 0, the Slater condition is satisfied, and there exists a point f such that F is continuous in f, for example $f(x) = 1 + h(x)$. For being able to apply the Kuhn-Tucker theorem, it remains ([6]) to show that $F(f)$ is closed. We will show that the set $\{F(f) < \tilde{\alpha}\}$ is open for all $\tilde{\alpha} \in \mathbb{R}$. Indeed, let firstly f be nonnegative. Then, by the definition of C, for $g \in C$ with $||g - f|| \leq t$,

$$g(x) \leq f(x) + t(1 + h(x))$$

Furthermore, as $t \downarrow 0$, $\alpha(x, f(x)+t(1+h(x))) \downarrow \alpha(x, f(x))$ almost surely, by the fact that α is nondecreasing and right-continuous in the second variable for $y \geq 0$. By FHFC, $\alpha(x, f(x) + t(1 + h(x))) \leq \alpha(x, h(x))$ which is $\nu-$ integrable. By the dominated convergence theorem, we must have that $F(f+t(1+h)) \downarrow F(f)$. For each $\epsilon > 0$, we can therefore find a $\delta > 0$ such that $F(g) \leq F(f+t(1+h)) < F(f) + \epsilon$ for all $||g - f|| \leq t < \delta$. Because $F(f) < \tilde{\alpha}$, we find an $\epsilon > 0$ such that $F(f) + \epsilon < \tilde{\alpha}$, and therefore $F(g) < \tilde{\alpha}$ for all g with $||f - g|| < \delta$. If f is not nonnegative, there exists a set $A' \subset A$, $\nu(A') > 0$, such that $f(x) < 0$ on A'. We may find a subset of A', again denoted by A', with $\nu(A') > 0$, such that $f(x) \leq -\epsilon < 0$. Because $h(x)$ is finite, we may furthermore find a further subset with nonzero $d\nu-$measure, again denoted by A', such that $h(x) \leq K < \infty$ for all $x \in A'$. Let now $||f - g|| < \frac{\epsilon}{2K+2}$. Then, on the set A',

$$g(x) \leq f(x) + \frac{\epsilon}{2K + 2}(1 + h(x)) \leq -\epsilon + \frac{\epsilon}{2} < 0$$

It follows that $\alpha(x, g(x)) = -\infty$ on a set with positive $d\nu-$ measure, and therefore $F(g) = -\infty < \tilde{\alpha}$ for all $\tilde{\alpha}$. As a consequence, $\{F(f) < \alpha\}$ is open, from which it follows that F is closed.
By the Kuhn-Tucker theorem in infinite dimensions [6], there must exist a continuous linear functional $\phi \in \delta F(v)$ in the superdifferential and a constant $\gamma > 0$ such that

$$\phi = \gamma\mu$$

By the fact that μ is absolutely continuous with respect to ν, it follows that

$$\phi(f) = \int_A \gamma \frac{d\mu}{d\nu}(x) f(x) d\nu(x)$$

that is ϕ is even in $L^1(A, \nu)$.
Now we define, for a function $f \in C$, the new function

$$g(t) := F(v + tf)$$

Let f be in C. Then $\phi(f)$ must be in the superdifferential of $\delta g(0)$, because we have

$$g(t) - g(0) = F(v + tf) - F(0) \leq \phi(f)(tf) = t\phi(f)$$

by the fact that ϕ is in the superdifferential of F. Let now f be such that g is continuous in 0 (that is $v + tf \geq 0$ for $|t|$ small enough). Then

$$g(t) - g(0) = \int_A [\alpha(x, v(x) + tf(x)) - \alpha(x, v(x))] \, d\nu(x) \leq t\phi(f) \quad \forall \quad t \in B_\epsilon(0)$$

must be satisfied. But for $t \downarrow 0$, we have

$$\lim_{t\downarrow 0} \frac{g(t)-g(0)}{t} = \lim_{t\downarrow 0} \int_A \frac{1}{t} \left[\alpha(x, v(x)+tf(x)) - \alpha(x,v(x))\right] d\nu(x)$$

The integrand converges $d\nu-$ a.s. to

$$\beta_-(x,v(x))f^+(x) - \beta_+(x,v(x))f^-(x)$$

where $\beta_-(x,y)$ is the right limit of the difference quotient of $\alpha(x,y)$ in y, and $\beta_+(x,y)$ the left limit, and f^+ and f^- are the nonnegative, respectively nonpositive, parts of f. Similarily, for $t \uparrow 0$, the integrand converges to

$$\beta_+(x,v(x))f^+(x) - \beta_-(x,v(x))f^-(x)$$

By the fact that $g(t)$ is concave and there exists an $\epsilon > 0$ with $g(-\epsilon) > -\infty$ the dominated convergence theorem yields

$$\begin{array}{rcl} \phi(f) & \in & \left[\int_A (\beta_-(x,v(x))f^+(x)d\nu(x) - \beta_+(x,v(x))f^-(x)d\nu(x)),\right. \\ & & \left.\int_A (\beta_+(x,v(x))f^+(x)d\nu(x) - \beta_-(x,v(x))f^-(x)d\nu(x))\right] \end{array} \quad (21)$$

Let now $\hat{\phi}(x) = \gamma \frac{d\mu}{d\nu}(x)$ be the function with $\phi(f) = \int_A \hat{\phi}(x)f(x)d\nu(x)$. Assume that on a set $A' \subset A$ with $v(x) > 0$ on A' and $\nu(A') > 0$, we have $\hat{\phi}(x) > \beta_+(x,v(x))$. Then we find a subset of A', denoted again by A', on which $v(x) \geq \epsilon > 0$, with $\nu(A') > 0$. The function $1_{A'}$ is obviously in C and nonnegative, and for this function, $g_{1_{A'}}(t)$ is continous at 0. It follows from (21) that

$$\phi(1_{A'}) \leq \int_{A'} \beta_+(x,v(x))d\nu(x)$$

But by definition,

$$\begin{array}{rcl} \phi(1_{A'}) & = & \int_{A'} \hat{\phi}(x)d\nu(x) = \int_{A'} \beta_+(x,v(x))d\nu(x) + \int_{A'} \left(\hat{\phi}(x) - \beta_+(x,v(x))\right) d\nu(x) \\ & > & \int_{A'} \beta_+(x,v(x))d\nu(x) \geq \phi(1_{A'}) \end{array}$$

a contradiction. In the same way, the assumption $\hat{\phi}(x) < \beta_-(x,v(x))$ leads to a contradiction on $\{v(x) > 0\}$.
On $\{v(x) = 0\}$, if $f(x) \geq 0$ but $g(t)$ is not necessarily continuous at 0, we can still take the right limit $g(t) \downarrow 0$, and the monotone convergence theorem yields

$$\phi(f) \geq \int_A \beta_-(x,v(x))f(x)d\nu(x)$$

Furthermore, we have $\beta_+(x,0) = \infty$ by the definition of α. We must have that $\beta_-(x,0) < \infty$, because otherwise, defining $f(x) = 1_{v(x)=0}(x)$, we would have $\phi(f) = \infty$, a contradiction to the continuity of ϕ. Let now $\hat{\phi}(x) < \beta_-(x,v(x))$ on a subset A' of $\{v(x) = 0\}$, with $\nu(A') > 0$. Again,

$$\begin{array}{rcl} \int_{A'} \hat{\phi}(x)d\nu(x) & = & \int_{A'} \beta_-(x,v(x))d\nu(x) - \int_{A'} \left(\beta_-(x,v(x)) - \hat{\phi}(x)\right) d\nu(x) \\ & < & \int_{A'} \beta_-(x,v(x))d\nu(x) \leq \phi(1_{A'}) \end{array}$$

which is the contradiction. It follows that always

$$\hat{\phi}(x) = \gamma \frac{d\mu}{d\nu}(x) \in [\beta_-(x, v(x)), \beta_+(x, v(x))]$$

For each $x \in A$, defining $y = v(x)$, we may therefore find a point $\beta(x, y)$ in this interval such that equation 18 holds. But by definition, this interval coincides precisely with the superdifferential of $\alpha(x, y)$ at point y. For a y for which no x exists with $v(x) = y$, we may choose an arbitrary point $\beta(x, y)$ in the superdifferential of α. It follows that the $\beta(x, y)$ defined in this way satisfies property 18, and the theorem is now completely proved. □

Remarks:

- For a continuous and strictly concave function, this theorem is proposition 5.14 in [3].
- If $\alpha(x, y)$ is strictly concave, Theorem 4.1 gives an algebraic equation where to find the optimal function $v(x)$.

Indeed the optimal function exists, which can be proved in the same way as in [3]. Theorem 4.1 is therefore just an answer to the question where an optimum may be found. Similar to [3], we state also the existence theorem.

Theorem 4.2. *Let μ and ν be two finite equivalent measures, and let $\alpha(x, y)$ be as in Theorem 4.1. Let $V_0 < \mu(h)$ be larger 0. Then there exists a measurable function $v(x) \in C$, C defined as in Theorem 4.1, such that*

$$\int_A \alpha(x, v(x)) d\nu(x) = \sup_f \int_A \alpha(x, f(x)) d\nu(x)$$

where the supremum is taken over all measurable functions f with

$$\mu(f) \leq V_0$$

Proof. Because $\alpha(x, y) = -\infty$ for $y < 0$, we can restrict to nonnegative functions f. Furthermore, we can restrict to functions in C, because $h(x) \in C$, and any nonnegative measurable function f satisfies

$$\int_A \alpha(x, f(x)) d\nu(x) = \int_A \alpha(x, f(x) \wedge h(x)) d\nu(x)$$

so that f can be even chosen bounded by $h(x)$. We can furthermore choose f such that $\mu(f) = V_0$, because if $\mu(f) < V_0 < \mu(h)$, $f_t(x) := f(x) + t(h(x) - f(x))$ is for $t \in [0, 1]$ still a nonnegative function bounded by h, and by the fact that $\alpha(x, y)$ is nondecreasing in y,

$$\int_A f_t(x) d\nu(x)$$

is a nondecreasing function in t, and

$$g(t) := \int_A f_t(x) d\mu(x)$$

is a continuous (linear) function with $g(0) = \mu(f)$ and $g(1) = \mu(h)$. By standard real analysis, there exists a $0 < t < 1$ with $\mu(f_t) = V_0$. If we define now the set

$$C' := \{0 \leq f(x) \leq h(x) : \mu(f) = V_0\}$$

C' is a convex set which is weakly compact in $\mathcal{L}^1$, and we are precisely in the same situation as in [3]. Existence follows now by the same arguments. □

Corollary 4.3. *Let $(\Omega, \mathcal{F}_T, \mathcal{F}_t, P)$ be a filtered probability space, and X_t a continuous semimartingale with values in the convex set A from before. Let there exist a unique equivalent local martingale measure Q, and assume that $\frac{dQ}{dP}$ is $\sigma(X_T)$-measurable. Let $\alpha(x, y)$ satisfy the properties of Theorem 4.1, with μ and ν the laws of X_T under Q and P, respectively. Then the hedge which optimizes $E[\alpha(X_T, V_T)]$ at initial capital V_0 is given by the hedge of the claim $v(X_T)$, with v from the Theorems 4.1 and 4.2.*

Proof. The proof follows similar arguments as given in [3]. Let π be any admissible strategy, and let its value process be

$$V_t = V_0 + \int_0^t \pi_s dX_s$$

At time T, we define the X_T-measurable random variable $f(X_T) := E[V_T|\sigma(X_T)]$. By the concavity of α in the second argument we have that

$$E[\alpha(X_T, V_T)] \leq E[\alpha(X_T, E[V_T|\sigma(X_T)])] = E[\alpha(X_T, f(X_T))]$$

But by the fact that $\frac{dQ}{dP}$ is $\sigma(X_T)$ measurable, we have

$$E^Q[f(X_T)] = E[\frac{dQ}{dP} E[V_T|\sigma(X_T)]] = E[E[\frac{dQ}{dP} V_T|\sigma(X_T)]] = E^Q[V_T] = V_0$$

If v is optimal in the sense of Theorem 4.1 or 4.2, it follows that

$$\begin{aligned} E[\alpha(X_T, V_T)] &\leq E[\alpha(X_T, f(X_T))] = \int_A \alpha(x, f(x)) d\nu(x) \\ &\leq \int_A \alpha(x, v(x)) d\nu(x) = E[\alpha(X_T, v(X_T))] \end{aligned}$$

and therefore the replication of the claim $v(X_T)$ is optimal. □

5 Application to the unit-linked insurance model

5.1 Structure of the optimal hedge

We take again the same model for the financial market as in section 3 as well as the same unit-linked survival insurance, but now with a discrete distribution of the amount of the survivors N_T which is not specified at this stage.
As in section 2.3, we are looking firstly for a strategy which solves the maximization problem (6) for fixed values of V_0 and α. As in section 3, information about the insurance process arrives only at time T, and the considerations of

section 2.4 are therefore applicable. Due to equation (7), the value function $\alpha(x, y)$, conditional $X_{T^-} = x$ and the fund value $V_{T^-} = y$, is

$$\alpha(x, y) = \sum_{j=0}^{n} p_j \left((jx - \alpha) \wedge y\right) \tag{22}$$

where p_j denotes the probability of j survivors. We wrote here $(jx-\alpha)$ instead of $(jx - \alpha)^+$, because from section 2.3 it is clear that we are interested in nonpositive values of α. ¡ As in section 2.4, we are now in a complete setting, with a unique equivalent martingale measure Q given by

$$\frac{dQ}{dP} = \gamma X_T^{-\frac{\mu}{\sigma^2}} \tag{23}$$

where $\gamma > 0$ is a constant.
If the amount of insured persons is finite, then

$$E[\alpha(X_T, Y_T)] \leq nE[X_T \wedge Y_T] - \alpha \leq nE[X_T] - \alpha$$

and therefore FHFC is satisfied. We will apply now Theorem 4.1. The superdifferential of α with respect to y is

$$\delta\alpha(x, y) = \begin{array}{ll} \sum_{j=k}^{n} p_j & \text{if} \quad 0 \vee ((k-1)x - \alpha) < y < kx - \alpha \\ \left[\sum_{j=k+1}^{n} p_j, \sum_{j=k}^{n} p_j\right] & \text{if} \quad y = kx - \alpha \end{array}$$

In order to apply Theorem 4.1, we want to choose a specific function $\beta(x, y)$ in the superdifferential with the property

$$\beta(X_T, v(X_T)) = \gamma X_T^{-\frac{\mu}{\sigma^2}}$$

Because the right-hand side has no constant points, it is reasonable to choose $v(x)$ in a way that for each x there exists a $k_x \in \mathbb{N}$ with $v(x) = k_x x - \alpha$, and we must have that for any $k \leq n$ that $v(x) = kx - \alpha$ implies

$$\gamma X_T^{-\frac{\mu}{\sigma^2}} \in [\sum_{j=k+1}^{n} p_j, \sum_{j=k}^{n} p_j]$$

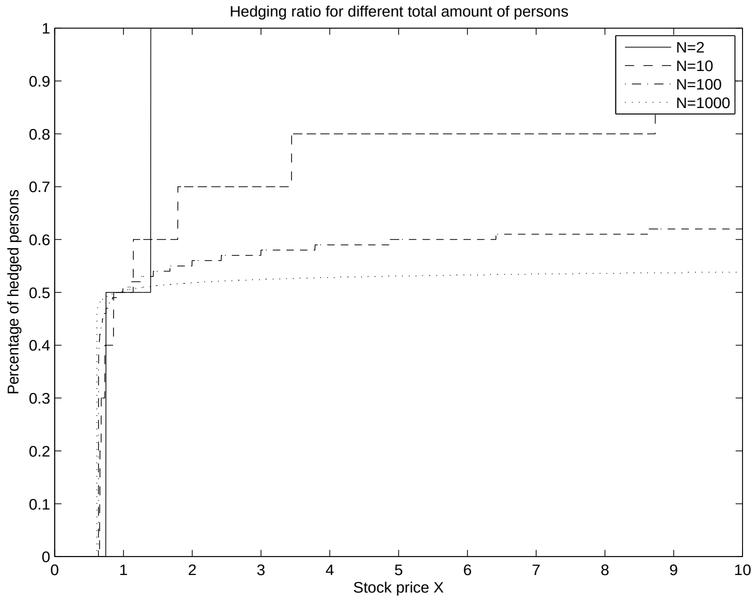
and therefore

$$v(x) = \begin{array}{lll} 0 & \text{if} & x < \gamma^{\frac{\sigma^2}{\mu}} \\ kx - \alpha & \text{if} & x \in [\left(\frac{\gamma}{\sum_{j=k}^{n} p_j}\right)^{\frac{\sigma^2}{\mu}}, \left(\frac{\gamma}{\sum_{j=k+1}^{n} p_j}\right)^{\frac{\sigma^2}{\mu}}] \\ nx - \alpha & \text{if} & x > \left(\frac{\gamma}{p_n}\right)^{\frac{\sigma^2}{\mu}} \end{array}$$

We have therefore proved the following theorem:

Theorem 5.1. *Let the financial market and the unit linked insurance product satisfy the assumptions from this section. Then there exist constants* $c_0 < c_1 <$

Figure 1: Optimal payoff for different amount of insured persons



$c_2 < .. < c_n$ such that the hedge which minimizes the expected shortfall is given by the hedge of the replicable claim

$$v(X_T) := \sum_{k=0}^{n-1} (kX_T - \alpha) 1_{c_k < X_T < c_{k+1}} + (nX_T - \alpha) 1_{X_T > c_n}$$

Proof. We only have to set

$$c_k := \left(\frac{\gamma}{\sum_{j=k}^{n} p_j} \right)^{\frac{\sigma^2}{\mu}} \tag{24}$$

Then $v(x)$ satisfies the assumptions of Theorem 4.1, and is therefore optimal. $\square$

Remark

- Even if under the original probability measure the processes N_T and X_t are independent, the optimal hedging strategy does not simply hedge npX_T, but hedges a higher survival rate for larger values of X_T. This means also that under the worst-case martingale measure as stated in [3] the two events are not independent any more.

As an illustration, figure 1 shows the hedge ratio of the optimal payoff to be hedged as a function of the final stock price, for different amount of insured persons. From this, one can see that the optimal payoff is a sum of knock-in options.
For only few persons (in practice, one can think about an special insurance with insurance sum which exceeds by far limit for mass business), zero payoff

is hedged if the stock price is below a limit c_1, whereas a full hedge is done if the stock price is large enough. The reason is that, if the stock price is large, the risk that there are more survivors than expected plays a much larger role than for small stock prices. If the amount of insured persons increases, the optimal hedge ratio converges more and more to the one which is usually done by actuaries in practice, namely the hedge of the expected amount of survivors. From figure 1, one can also see one important risk when performing such an optimization in expected shortfall: For experienced actuaries, the constant c_1 below which zero payoff is hedged seems too large for an amount of 1000 insured persons. The reason is that under the assumption of lognormal stock returns (geometric Brownian motion), the probability that the stock price goes below the level c_1 is underestimated with respect to the reality, because the kurtosis in most of the stock prices is larger than under this assumption. The optimization has only been only done under the assumption of a specific model. This optimal payoff depends strongly on the underlying model.
For practical application, it is essential that not only the risk due to incompleteness, but also the model risk is considered. It is necessary to make stress tests of the optimal payoff function with respect to different model assumptions.
On the other hand, this shows also that in general the risk depends strongly on the underlying model assumptions.

5.2 Pricing with the CVaR criterion

We will now calculate the minimal capital V_0 such that the claim becomes acceptable due to criterion (1). For this, we will assume that the amount of survivors is binomially distributed. As in equation (5), we are looking firstly at the minimal expected shortfall at a given α and initial capital V_0. From Theorem 5.1, the optimal payoff for a fixed α is

$$\sum_{k=0}^{n}(kX_T-\alpha)1_{c_{k+1}>X_T>c_k}$$

with c_k as in equation (24) with $c_{n+1}=\infty$. It follows that, for fixed α,

$$V_0(\alpha)=\sum_{k=1}^{n}X_0\Phi\left(\frac{\ln\left(\frac{X_0}{c_k}\right)+\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right)-\alpha\Phi\left(\frac{\ln\left(\frac{X_0}{c_0}\right)-\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) \tag{25}$$

where Φ is the cumulated normal distribution, and c_k are determined by (24) and the constraint

$$\begin{aligned}-\alpha(1-\beta) \quad &= \quad X_0e^{\mu T}\textstyle\sum_{k=1}^{n}\left(\sum_{l=k}^{n}p_l\right)\Phi\left(\frac{\ln\left(\frac{c_k}{X_0}\right)-(\mu+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right)\\ &\quad -\alpha\Phi\left(\frac{\ln\left(\frac{c_0}{X_0}\right)-(\mu-\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right)\end{aligned} \tag{26}$$

We set here the equality sign because we are here precisely in the comfortable situation where $f_{min}(\alpha,V_0)$ is continuous in V_0.

Table 2: Minimal capital for different parameters

n	p	γ	α	V_0	*Load*
1000	0.5	0.052	-6.72	532.60	6.5%
1000	0.1	0.052	-4.24	120.0	20.0%
50	0.5	0.052	-1.48	32.24	29.0%
50	0.1	0.052	-1.09	9.76	95.2%

In equation (26) we can trivially find for each value of γ from (24) the corresponding value of α, and set this into (25). That is we do, as in section 3, a minimization of (25) in γ.
Even if the CVaR is a translation invariant risk measure, it is not the same to calculate only the minimal CVaR at capital 0 and take this as the minimal capital required. The reason is that for larger initial capitals, more trading strategies are allowed.
We repeat the numerical example from section 3, with the same parameters as there. The results are in table 5.2. Comparing the results with the ones of section 3, one can see that the differences to the normal approximation are rather small, even for the case of only 50 persons.

6 General information structure

6.1 Information about amount of survivors during hedging process

Usually, it is not realistic to assume that there is no information about the insurance process until time T. Usually the process for the insurance risk follows a continuous-time process Z_t with terminal value Z_T. As a first step, there is a time t_0 with $0 < t_0 < T$ where the information about the insurance process Z_{t_0} arrives. This information can therefore be used for the further hedge.
Let us be already at time t_0 such that we know the variable X_{t_0}, the capital V_{t_0} and the insurance outcome Z_{t_0}. Then our problem is precisely the one of section 2.4, and the aim is to optimize the expectation

$$E[h(X_T, Z_T)|\mathcal{F}_{t_0}]$$

under the restriction that the initial value V_{t_0} of the wealth process is not larger than V_{max}, where $h(x, y)$ is the measurable function defined in (7) which is concave in y, with exception that now the function h may depend on the outcome Z_{t_0}, because the distribution function $dP(Z_T)$ in the integral (7) is now conditional on Z_{t_0}. By the Markov property of X_t, we have, given Z_{t_0}, that the unique density process $\frac{dQ}{dP}$ is a function of X_T, and we may therefore apply Corollary 4.3 to get a measurable function $v_{V_{max}}$ such that $v_{V_{max}}(X_T)$ is the optimal claim to hedge given the information at time t_0. The optimal value at time t_0 is therefore

$$u(X_{t_0}, V_{t_0}, z) = E[h(X_T, v_{V_{t_0}}(X_T))|Z_{t_0} = z] \tag{27}$$

where z denotes the insurance outcome at time t_0.

Proposition 6.1. *For all z, the function u defined in equation (27) is nondecreasing and concave concave in the second argument is the function h is.*

Proof. For simplifying the notion, we set $X_0 := X_{t_0}$ and $V_0 := V_{t_0}$. Let V_0 and $\tilde{V}_0$ be two initial capitals. Nondecreasing is trivial because if $\tilde{V}_0 > V_0$, we have

$$\begin{aligned} u(X_0, V_0, z) &= \sup_\pi E[h(X_T, Z_T^\pi)|X_0, Y_0 \leq V_0, Z_{t_0} = z] \\ &\leq \sup_\pi E[h(X_T, Y_T^\pi)|X_0, Y_0 \leq \tilde{V}_0, Z_{t_0} = z] = u(X_0, \tilde{V}_0, z) \end{aligned}$$

We have by the concavity of h that

$$\begin{aligned} tu(X_0, V_0, z) + (1-t)u(X_0, \tilde{V}_0, z) &= E[th(X_T, v_{V_0}(X_T)) + (1-t)h(X_T, v_{\tilde{V}_0}(X_T))|\mathcal{F}_{t_0}] \\ &\leq E[h(X_T, tv_{V_0}(X_T) + (1-t)v_{\tilde{V}_0}(X_T))|\mathcal{F}_{t_0}] \end{aligned} \tag{28}$$

But by definition

$$B_{t_0} E^Q[v_{V_0}(X_T)|\mathcal{F}_{t_0}] \leq V_0$$

where B_t is the price process of the zero bond with expiry at time T. The same holds for $\tilde{V}_0$, so that the claim $tv_{V_0}(X_T) + (1-t)v_{\tilde{V}_0}(X_T)$ has under Q the expectation of less or equal $tV_0 + (1-t)\tilde{V}_0$. This claim is therefore (super-) replicable at initial capital $tV_0 + (1-t)\tilde{V}_0$, and the optimal value available at this capital must satisfy

$$E[h(X_T, tv_{V_0}(X_T) + (1-t)v_{\tilde{V}_0}(X_T))|\mathcal{F}_{t_0}] \leq E[h(X_T, v_{tV_0+(1-t)\tilde{V}_0}(X_T))|\mathcal{F}_{t_0}]$$

where the function $v_{tV_0+(1-t)\tilde{V}_0}(X_T)$ denotes the optimal claim available at capital $tV_0 + (1-t)\tilde{V}_0$ Together with equation (28), this implies the concavity of the function u. □

At time t_0^-, we do not know yet the outcome Z_{t_0} at time t_0, and our value function is therefore

$$h_0(X_{t_0}, V_{t_0}) = \int u(X_{t_0}, V_{t_0}, z) dP_{Z_0}(z) \tag{29}$$

where dP_{Z_0} denotes the distribution function of Z_0. This is a convex combination of concave functions and therefore still concave. As a consequence, we can repeat our procedure for arbitrary many $t_n < t_{n-1} < ..t_1 < t_0 < T$.
As all traded assets, B_{t_0} is assumed to be a function of X_{t_0}, so that a stochastic interest rate can be included easily.

6.2 Limiting case: Continuous-time information

The aim of this subsection is to prove the following theorem:

Theorem 6.2. *Let the situation be as before, but let now Z_t, the insurance process, be a continuous-time process. Let the claim $g(X_T, Z_T)$ from (6 be integrable. Let Π_k be a sequence of partitions of $[0,T]$, and let $u_k(x,y)$ be the optimal solution using the method above, with information about mortality at* *the points* $0 \le t_1^{(k)} < .. < t_k^k \le T$*, that is at the points of the partition, and let* $u(x,y)$ *be the optimal value function in continuous-time. Then, if the mesh of* Π_k *converges to 0 as* $k \to \infty$ *and* Π_{k+1} *is always a refinement of* Π_k*, we have for the value functions that* $u_k(x,y) \uparrow u(x,y)$*.*

The proof is divided in different lemmas:

Lemma 6.3. *If* Π_{k+1} *is a refinement of* Π_k*, then* $u_{k+1}(x,y) \ge u_k(x,y)$*.*

Proof. The strategy leading to the value $u_k(x,y)$ is allowed also if information at the time points Π_{k+1} takes place (by only considering the information at time points Π_k). Because $u_{k+1}(x,y)$ is the optimal value for the information times Π_{k+1}, it must be larger or equal $u_k(x,y)$. □

Proposition 6.4. *If* $g(X_T, Z_T)$ *is integrable, the sequence* $u_k(x,y)$ *converges pointwise to a limit.*

Proof. By Lemma 6.3, for each fixed point (x,y), $u_k(x,y)$ is a monotonically increasing sequence of real numbers. Furthermore

$$u_k(x,y) \le E[(g(X_T, Z_T) - \alpha)^+ | X_0 = x, V_0 = y]$$

and therefore, for all (x,y), the sequence is bounded from above. Pointwise convergence is now standard real analysis. □

Proposition 6.5. *Let the claim* $g(X_T, Z_T)$ *to be hedged be integrable, and* π^* *be the optimal strategy at initial capital* y *and initial state* $X_0 = x$*. Assume that this strategy is càglàd, and the trated assets* S *in the financial market are continuous. Then, for every sequence of partitions* Π_k *with mesh converging to 0, value functions* $\tilde{u}_k(x,y)$ *of the constant policy strategy*

$$\tilde{\pi}_k(t) := \begin{array}{ll} \pi_k(t) & \text{if} \quad t < \hat{T}^{(k)} \\ 0 & \text{if} \quad t \ge \hat{T}^{(k)} \end{array} \tag{30}$$

with

$$\pi_k := \sum_{j=1}^{k} \pi^*(T_{j-1}^{(k)}) 1_{(T_{j-1}^{(k)}, T_j^{(k)}]} \tag{31}$$

converge pointwise to the optimal value function $u(x,y)$*, where* $T_j^{(k)}$ *is the sequence of stopping times defined inductively in* j *by* $T_0^{(k)} = t_0$ *and*

$$T_j^{(k)} := \begin{array}{ll} t_j & \text{if} \quad \tilde{Q}_{j-1,k}(t) > 0 \quad \forall \quad t \le t_j \\ t_j & \text{if} \quad \exists t \le T_{j-1}^{(k)} : Q_{j-1,k}(t) \le 0 \\ \inf\{t > 0 : \tilde{Q}_{j-1,k}(t) \le 0\} & \text{otherwise} \end{array} \tag{32}$$

where Q_{jk} *and* $\tilde{Q}_{jk}$ *are defined again inductively by*

$$\begin{array}{rcl} Q_{jk}(t) & := & V_0 + \sum_{l=1}^{j} \pi^*(T_{l-1}^{(k)}) \left(S_{t \wedge T_l^{(k)}} - S_{t \wedge T_{l-1}^{(k)}} \right) \\ \tilde{Q}_{jk}(t) & := & Q_{jk}(t) + \pi^*(T_j^{(k)}) \left(S_t - S_{t \wedge T_j^{(k)}} \right) \end{array}$$

and $\hat{T}^{(k)}$ is the stopping time

$$\hat{T}^{(k)} := \inf\left\{t > 0 : V_0 + \int_0^t \pi_k dS \leq 0\right\} \tag{33}$$

Proof. Let V_T^* the optimal capital from strategy π^*, that is

$$V_T^* = V_0 + \int_0^T \pi_t^* dX_t$$

We have the obvious equality that

$$\int_0^T \tilde{\pi}_k dS = \int_0^T \pi_k dS + \int_0^T (\tilde{\pi}_k - \pi_k) dS$$

By standard theory ([11]), the random variables V_k, defined by

$$V_k := V_0 + \int_0^T \pi_k dS = V_0 + \sum_{j=1}^k \pi^*_{T_{j-1}}(S_{T_j} - S_{T_{j-1}})$$

converge in probability to the random variable V_T^*. For showing that the second term converges to 0 in probability, it is, by the continuity of the integral operator [11], enough to show that the integrand converges in *ucp* to 0, because it is a sequence of simple predictable processes.
Let $A := \{V_T^* > 0\}$. Then, for each fixed $\omega \in A$, $V_t^*(\omega) > 0$ for all $t \leq T$, because $V_t^* = 0$ on an $\mathcal{F}_t$-measurable subset of A and $V_T^* > 0$ on A would be an arbitrage opportunity. By the continuity of S_t, the continuity of V_t^* follows, and therefore the uniform continuity on $[0, T]$. It follows that A can be written as

$$A = \cup_{n \geq 1} A_n := \cup_{n \geq 1}\left\{\omega : \inf_t V_t^*(\omega) \geq \frac{1}{n}\right\}$$

We have

$$P[\{|\tilde{\pi}_k - \pi_k| > \epsilon\} \cap A] \leq P[\cup_{n \geq N} A_n] + P[\{|\tilde{\pi}_k - \pi_k| > \epsilon\} \cap A_N]$$

and

$$P[\{|\tilde{\pi}_k - \pi_k| > \epsilon\} \cap A_N] \leq P[\{V_0 + \inf_t \int_0^t \pi_k(r) dS_r \leq 0\} \cap A_N]$$

But for fixed N, the right-hand side converges to 0, because of the *ucp* convergence of the integrand to V_t^*, and the fact that $V_t^* \geq \frac{1}{N}$ uniformly on A_N. Therefore, for each $\delta > 0$, we can choose N large so that $A \setminus A_N$ has probability smaller $\frac{\delta}{2}$, and then choose k large so that the other term becomes small. We have therefore *ucp* convergence of $\tilde{\pi}_k - \pi_k$ to 0 on A, and so does the integral. On A^c, that is on $\{V_T^* = 0\}$, we have either $\tilde{\pi}_k = \pi_k$ or $V_0 + \int_0^T \tilde{\pi}_k dS = 0$, by the definition of $\tilde{\pi}_k$. In both cases, the stochastic integral converges to 0. It follows that also $\tilde{\pi}_k$ is an approximating sequence of strategies. By definition, those strategies are nonnegative.
It follows that the sequence $(g(X_T, Z_T) - \alpha)^+ \wedge V_k)$ converges in probability

to $\alpha(X_T, V_T^*)$. But because $(g(X_T, Z_T) - \alpha)^+$ is integrable, the sequence is dominated by an integrable random variable. It follows that the sequence is uniformely integrable. It converges therefore in L^1 to $(g(X_T, Z_T) - \alpha)^+ \wedge V_T^*)$, and therefore

$$\tilde{u}_k(x, y) \to u(x, y)$$

□

Proposition 6.6. *If there does not exist an optimal strategy which is càglàd, there exists a subsequence of partitions Π_{k_j} such that Proposition 6.5 does still hold for this subsequence.*

Proof. There exists a sequence of strategies, π_l, say, such that

$$\hat{u}_l(x, y) := E[(g(X_T, Z_T) - \alpha)^+ \wedge (V_0 + \int_0^T \pi_l dS_t)|X_0 = x, V_0 = y] \to u(x, y)$$

pointwise in (x, y). For each $\hat{u}_l$, we may apply the same arguments as in Proposition 6.5, and we obtain, for the partitions Π_k, a constant policy approximation $\tilde{u}_{kl}(x, y) \to \hat{u}_l(x, y)$. Let now $\epsilon > 0$. Then, for each fixed (x, y) and for each $j \geq 1$ we find an l_j with $|\hat{u}_{l_j}(x, y) - u(x, y)| < 2^{-j}$. Similarily, we find a k_j with $|\tilde{u}_{k_j l_j}(x, y) - \hat{u}_{l_j}(x, y)| < 2^{-j}$ for each j. The sequence $w_j(x, y) := \tilde{u}_{k_j l_j}(x, y)$ then converges pointwise to $u(x, y)$, and is of the form (31) for the partitions Π_{k_j}. □

We apply now Proposition 6.6 to the case where the optimal strategy is predictable but not necessarily left-continuous.

Lemma 6.7. *Let π be a predictable $S-$ integrable process. Then there exists a sequence π_n of càglàd processes such that*

$$\int_0^T \pi_n dS \to \int_0^T \pi dS \tag{34}$$

almost surely.

The result seems to be quite standard, but we did not find a suitable reference. We will give therefore a proof in the appendix.

Corollary 6.8. *If the optimal strategy π^* from Proposition 6.5 is predictable and $S-$ integrable but not necessarily càglàd, Proposition 6.5 does still hold, provided the claim $g(X_T, Z_T)$ to be hedged is integrable.*

Proof. From Lemma 6.7, there exists a strategy π_n satisfying 34. The sequence

$$(g(X_T, Z_T) - \alpha)^+ \wedge Y_T^{\pi_n} \to (g(X_T, Z_T) - \alpha)^+ \wedge Y_T^{\pi*}$$

almost surely and is dominated by $(g(X_T, Z_T) - \alpha)^+$. It follows that the value supremum over all càglàd strategies is the same as the value of the optimal predictable strategy, and the result follows from Proposition 6.6. □

Now we turn to the proof of Theorem 6.2:

Theorem 6.2. By Lemma 6.3 and Proposition 6.4, we have already seen that $u_k(x,y)$ converges monotonically to a limit $\bar{u}(x,y)$. By the optimality of $u(x,y)$, we must have $\bar{u}(x,y) \leq u(x,y)$. Let u_k now be fixed, and Π_k be the partition for it. Using this partition, we may define a constant policy $\tilde{\pi}_k$ as in equation (31). This strategy uses only the information about Z_t at times $t_j^{(k)}$ and is therefore allowed also in the model which uses information only at times in Π_k. By the optimality of $u_k(x,y)$ for this information, it follows that

$$u_k(x,y) \geq \tilde{u}_k(x,y)$$

By Proposition 6.6, there exists a subsequence of partitions Π_{k_j} such that $\tilde{u}_{k_j}(x,y) \to u(x,y)$, and therefore

$$u(x,y) \geq \bar{u}(x,y) \geq u_{k_j}(x,y) \geq \tilde{u}_{k_j}(x,y) \to u(x,y)$$

By the monotonicity of $u_k(x,y)$ in k, the result follows. □

6.3 Limit for CVaR

The aim of this section is to apply the continuous time limit results to the CVaR criterion. For this, we reconsider the considerations of section 2.3 in order to see that the function f_{min} there is determined here by

$$f_{min}^{(n)}(\alpha, V_0) = E\left[(g(X_T, Z_T) - \alpha)^+\right] - u_n(X_0, V_0; \alpha)$$

and the same for f_{min} with u instead of u_n. Let V_{max} be the minimal initial capital such that a superhedge of $(g(X_T,Z_T) - \alpha)^+$ in the continuous information model is possible.

Proposition 6.9. *We have the following:*

1. *The functions $f_{min}^{(n)}$ and f_{min} are strictly decreasing and strictly convex in V_0, are 0 at V_{max} and $E\left[(g(X_T,Z_T) - \alpha)^+\right]$ at $V_0 = 0$.*

2. *It follows that the functions are continuous as well as invertible in this interval. Furthermore, $(f_{min}^{(n)})^{-1} \to f_{min}^{-1}$ pointwise.*

3. *We have $V_0^{(n)}(\alpha) = (f_{min}^{(n)})^{-1}(\alpha, -\alpha(1-\beta))$ and $V_0(\alpha) = f_{min}^{-1}(\alpha, -\alpha(1-\beta))$, where $V_0^{(n)}(\alpha)$ is the minimal capital such that there exists a strategy satisfying (3) for information at time points of partition Π_n and $V_0(\alpha)$ the one for the continuous-time information.*

4. *If α^* minimizes $V_0(\alpha)$, then $V_0(\alpha^*)$ is the price of the claim for continuous-time information due to the CVaR criterion.*

5. *If α_n^* minimizes $V_0^{(n)}(\alpha)$, then $V_0(\alpha^*)$ converges monotonically decreasing to $V_0(\alpha^*)$.*

Proof. The essential statement is statement 1. Everything else is real analysis or follows by our definitions.
It is clear that $f_{min}^{(n)}$ and f_{min} are nonincreasing and convex, because u_n and u are nondecreasing and concave. Furthermore $u_n(x, 0; \alpha) = 0$ and the same for u, because with 0 wealth one cannot go into any hedging position different from 0 without the possibility that the wealth becomes negative. It follows that $f_{min}^{(n)}(\alpha, 0) = E\left[(g(X_T, Z_T) - \alpha)^+\right]$ and the same for f_{min}. At V_{max}, we have a superhedge, and therefore $f_{min}(\alpha, V_{max}) = 0$. Because Z_T and X_T are independent and all traded assets are measurable with respect to the X_t process, the supremum $E^Q\left[(g(X_T, Z_T) - \alpha)^+\right]$ over all equivalent martingale measures Q is the same no matter whether the information about Z_t arrives before or at time T. By the optional decomposition theorem, this means that there exists also a superhedge at V_{max} when the information arrives only at time of the partition Π_n.
The fact that the functions are strictly decreasing and strictly convex follows by the fact that for all initial capitals V with $0 < V < V_{max}$, the functions must be strictly between 0 and $= E\left[(g(X_T, Z_T) - \alpha)^+\right]$, and the fact that they are nonincreasing and convex.
The second statement follows from the first, and the result from real analysis that if we have $f_n \to f$, f_n and f decreasing and invertible, then $f_n^{-1} \to f^{-1}$. The other statements follow from what is stated before. □

7 Numerical issues

As a numerical illustration, we will discuss the solution of the problem treated in section 5 for the case of two insured persons, and compare different numerical solution procedures.

7.1 Analytical solution with only information at the end

By theorem 5.1, we have almost an analytic solution for the value function. Indeed, for the case of two insured persons, there exist constants c_1 and c_2 such that the optimal payoff is given by

$$v(X_T) = X_T 1_{X_T > c_1} + X_T 1_{X_T > c_2}$$

with $c_1 < c_2$ determined by (24) and the budget constraint. This is the sum of two knock-in options, and therefore, by taking the expectations,

$$E^P[v(X_T) \wedge N_T X_T] = e^{\mu T} X_0 \left((p_1 + p_2)\Phi\left(\frac{\ln\left(\frac{X_0}{c_1}\right) + (\mu + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) + p_2 \Phi\left(\frac{\ln\left(\frac{X_0}{c_2}\right) + (\mu + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right)\right) \tag{35}$$

with the price

$$E^Q[v(X_T)] = X_0 \Phi\left(\frac{\ln\left(\frac{X_0}{c_1}\right) + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) + X_0 \Phi\left(\frac{\ln\left(\frac{X_0}{c_2}\right) + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) \tag{36}$$

With an initial capital V_0, the constant γ in (24) and therefore the C_i can be determined numerically by solving equation (36) which is monotonically decreasing in γ.
The corresponding hedging strategy is then the delta of equation (36).
In the sequel, we will use this almost analytical solution to compare our numerical results from the following sections.

7.2 Solution by HJB equation

Let

$$V(t,x,y,n) := \inf_{\pi} E^P \left[(N_T X_T - V_0 - \int_0^T \pi dX)^+ | X_t = x, V_0 + \int_0^t \pi dX = y, N_t = n \right] \tag{37}$$

be the value function. By the Hamilton-Jacobi-Bellman equation and the Itô formula for semimartingales [11] it follows that the value function satisfies the following partial differential equation

$$\frac{\partial V}{\partial t} = \inf_{\pi} \mathcal{L} V \tag{38}$$

where the spacial differential operator $\mathcal{L}$ is given by

$$\begin{aligned} \mathcal{L}V \;\; = \;\; & \left(\mu x \tfrac{\partial}{\partial x} + \tfrac{\sigma^2}{2} x^2 \tfrac{\partial^2}{\partial x^2} + \pi \mu x \tfrac{\partial}{\partial y} + \tfrac{1}{2} \pi^2 \sigma^2 x^2 \tfrac{\partial^2}{\partial y^2} + \pi \sigma^2 x^2 \tfrac{\partial^2}{\partial x \partial y} \right) V \\ & + \lambda n (V(n-1) - V(n)) \end{aligned} \tag{39}$$

where $\lambda = -\ln(1-p)$ is the density where persons die, and $V(n) = 0$ for $n \leq 0$. We have calculated the value function using a constant-policy explicit approximation as well as an iterative implicit method proposed in [13]. In the case of information only at the end of the period, both methods give quite good approximations of the value function. However, the hedge deviates from the expected one by the analytical solution. It seems that the nondifferentiability in the terminal condition (37) has more impact here than in simple solution of a PDE without optimization procedure.

7.3 Lower bounds with Fast Legendre Transformation

In the case of only information at the end, we have, by equation (18) a way how to find the optimal payoff function. Indeed, because we know that the Radon-Nikodym density is proportional to $X_T^{-\frac{\mu}{\sigma^2}}$, we have that

$$v(x) = \frac{\partial}{\partial y} \alpha^*(x, \gamma x^{-\frac{\mu}{\sigma^2}}) \tag{40}$$

where α^* is the Legendre transform of α from Theorem 4.1, and γ is a constant which has again to be determined iteratively by the budget constraint. The derivative is not unique because α^* has some points where it is not differentiable, but the problem appears only on a nullset and can therefore be thought as negligible.

A better lower bound can be obtained applying more information steps, as described in section 6.1. This can be done easily by making the one-step procedure more than once, always conditioning on the information about the insurance process which is already available.
The good news is that the calculation of $\alpha_y^*(x, y; n)$ has only to be done once per step of calculation, on a grid of $(x, y; n)$, but not during the iteration for γ. This makes this inversion numerically tractable.
The largest problem is that we need still a grid of $(x, y; n)$ for calculating the values. However, the grid for the state variables is used only for calculating the expectations under the measure P and Q. This could also be done by a simulation procedure, or using an irregular grid. Therefore, this method could also be applied for problems in higher dimensions.
The method gives only the payoffs to be hedged. The hedging strategy itself has then to be determined by the usual delta, applied to payoff determined by the Fast Legendre method.

7.4 Numerical experiments

In Figure 2, we plot the value function, for information only at the end of the period, for the case of two insured persons, and the parameter values $\mu = 0.07$, $\sigma = 0.2$, $T = 1$, $X_0 = 1$, $V_0 = X_0$ and the amount of survivors follows a binomial distribution with probability $p = 0.5$. In Figure 3, we plot the differences of the two numerical methods with respect to the analytical solution. One can see on the graphics that the value functions do not differ a lot from the analytic solution, neither the one calculated by HJB nor the one by Fast Legendre transformation.
In Figure 4, we show the differences in the value function when the amount of information points increases. This has been done by Fast Legendre transformation method. It may be a surprising issue that the value function cannot be improved a lot when adding more information. On the other hand, the error due to the space discretization in the Fast Legendre tranformation method increases when adding more information points. This problem can only be solved by choosing a finer space discretization.

Finally, one is also interested in the hedging position to be taken for the optimal hedge. We compare, for the parameters the same as before, the analytic delta hedge with the positions calculated by the HJB method. Figures 5 and 6 compare the delta-hedging positions for the same parameters as above, if the current wealth is $V_0 = 1$, for the analytical solution and the one obtained by HJB.
From Figures 5 and 6, one can see firstly that there is still an error of sometimes more than 10% in the HJB method, even if the computational effort was already rather large. The results depend also on the question which stencil has been taken for the mixed derivatives. Further research may be put to the question if an improvement can be obtained when going to an irregular grid, because in this way the mesh can be made fine precisely at the positions where it is needed.

Figure 2: Optimal value function

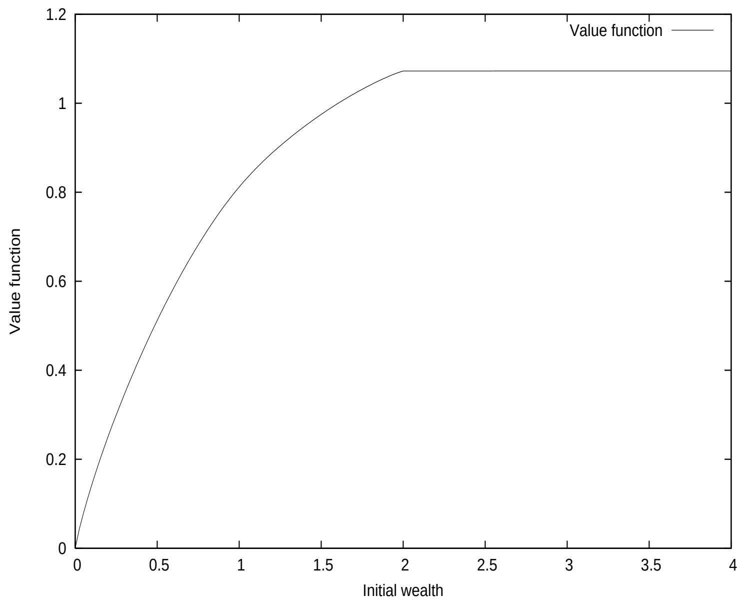


Figure 3: Comparison of the value functions among different methods

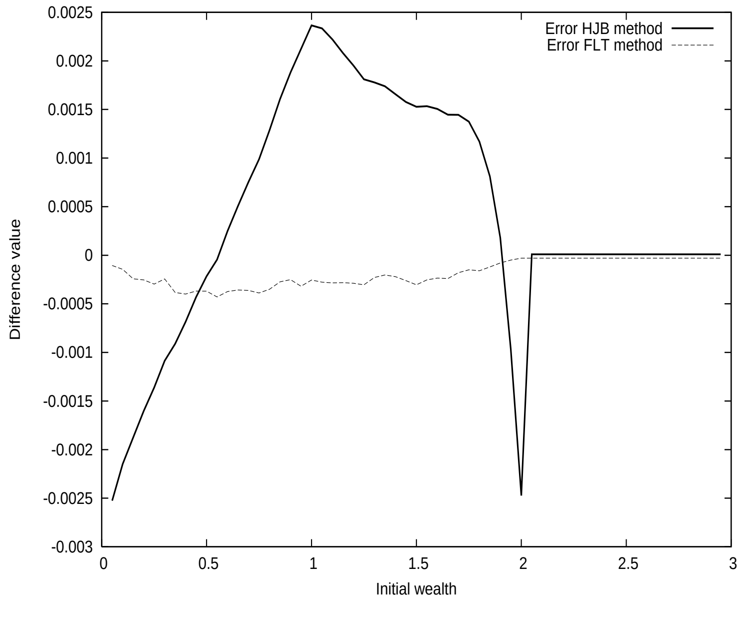


Figure 4: Comparison of the value functions with increasing information

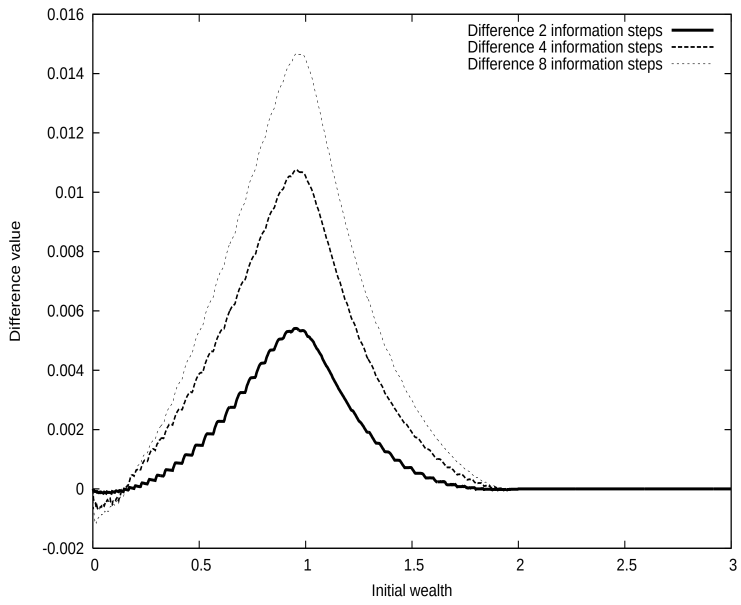


Figure 5: Hedging positions for time horizon T=0.5

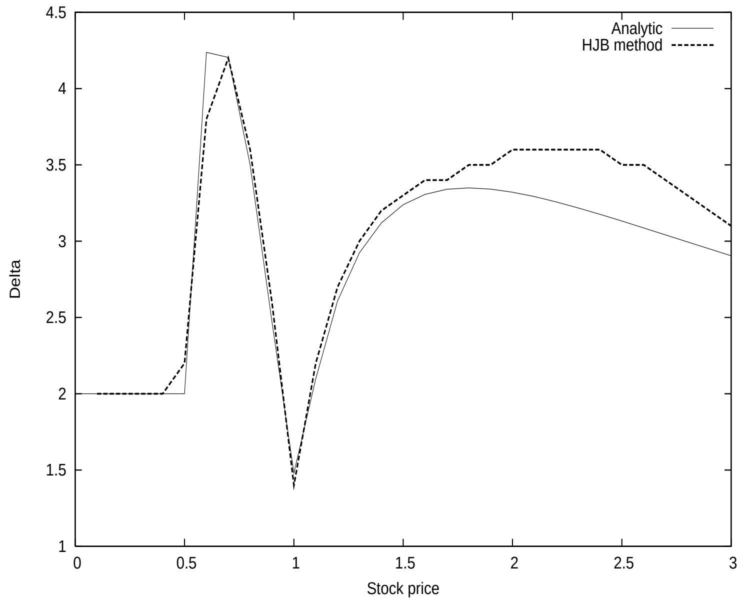
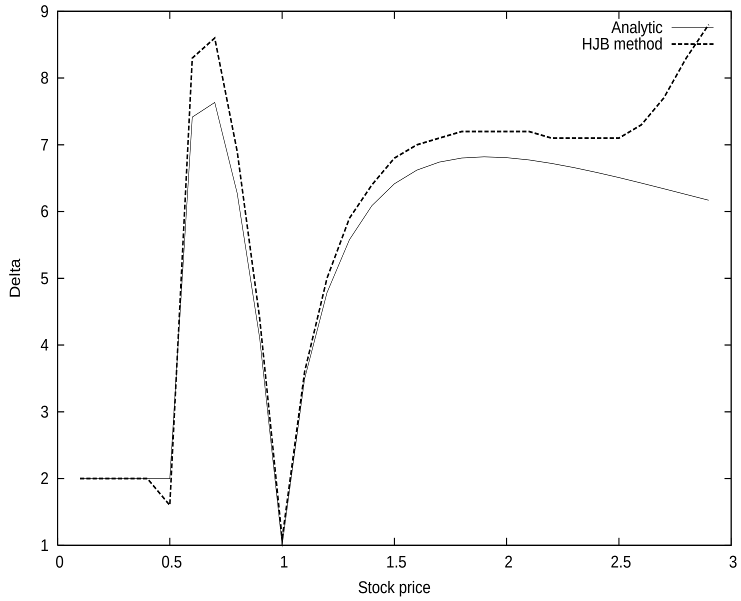


Figure 6: Hedging positions for time horizon T=0.1



Second, one can also see what has already been looked at in [14]: For small time horizons, the delta positions become very large. Actually, for being able to solve the HJB equation, we had to fix a maximal delta position in order to be able to make the computation. However, this is not the reason for the deviation of the numerical method from the HJB solution.

8 Conclusion

In the context of unit-linked insurance products, we have seen how upper bounds for the minimal expected shortfall as well as the corresponding hedge can be found by means of a discretization of the information flow of the insurance process. We proved furthermore the convergence of those bounds to the minimal expected shortfall as the mesh of this time discretization goes to 0.
We followed for this the approach of Föllmer/Leukert, which connects this problem to the Neyman-Pearson theory. However, we had to prove an extension and adjustment of this theory. In order to find the optimal payoff to be hedged, we have seen that the utility function does not need to be differentiable nor strictly concave in order to obtain that the optimal payoff is a generalization of the Legendre transform of the utilty function. This result allows us to find the optimal hedge in the case where the utility function is piecewise linear, as is typically the case when dealing with CVaR and the possibility of jumps at the terminal time in the insurance procss. As a further issue, we did not assume the existence of a risk-free interest rate; the existence of suitable zero bonds are enough.

In the specific case of a unit-linked survival insurance where information arrives only at the end of the period, the general results allow to obtain analytical solutions for the problem. We have seen that the optimal payoff function is a sum of knock-in options. It can be interpreted as a hedge of the expectation plus a hedge of the risk-loading. In contrast to what actuaries do traditionally, the risk loading should be hedged in a dynamic way in order to minimize the expected shortfall.
For this specific situation, we gave also an explicit formula for the worst-case martingale measure, in the sense of Neyman-Pearson. It has been shown that, even if the financial market and insurance process are independent under the original measure, they are not any more under this worst-case measure, which can also be interpreted as a risk-adjusted pricing measure.
The general results allow also to calculate upper bounds for the minimal expected shortfall numerically by Fast Legendre Transformation, in the case where no analytic solutions can be found. This provides an alternative to the classical Hamilton-Jacobi-Bellman approach. Even if the computational effort increases when calculating better bounds, one can see in our example that the first bound which can be calculated easily is already quite good and the corrections for better bounds are small.
Due to the way how risk measures are defined, we had to evaluate the risk always at a specific point in time, the terminal time of the contract. However, in insurance practice, we have typically payment processes where the payment takes place at different points in time. For further research, it may be interesting to look in more detail at this issue.
Another issue for further research are other applications than the hedging of a unit linked survival insurance. In particular, the case of a partial hedge of the longevity risk seems to be interesting for insurance practice.
A further extension would be to have more sophisticated models for the financial market, which we assumed in this paper to be driven by Brownian motions. This would violate one key assumption of this paper, namely that the source of incompleteness of the market is only the insurance process.
Last, the numerical computation remains a problem, especially for higher dimensional problems with stochastic interest rates, more assets and stochastic volatility models. We have already indicated that using the Fast Legendre transformation method, we could also use others than regular grids. To further develope this in order to obtain lower bounds in high dimensions may be an interesting topic.

A Proof of Lemma 6.7

Proof. If S is square integrable and π is bounded, it follows by Theorem 2 and 3 of chapter 4 in [11] that there exists a sequence of càglàd strategies π_n such that the integral $\pi_n \cdot S$ converges to $\pi \cdot S$ in the space of square-integrable semimartingales. If π is not bounded, we can approximate it by $\pi_m := \pi 1_{|\pi| \leq m}$ which are bounded, and the integral $\pi_m \cdot S$ converges to $\pi \cdot S$, by Theorem 14 of the same chapter, in the space of square-integrable semimartingales. It follows that there exists a subsequence of bounded càglàd processes such that the integral converges in this space. For a fixed $T > 0$, the sequence

$$\int_0^T \pi_{n_k} dS \to \int_0^T \pi dS \tag{41}$$

in probability, with π_{n_k} the subsequence from before. We have therefore a subsequence of π_{n_k}, denoted again π_{n_k}, such that the convergence is almost surely.

By definition in [11], a predictable process π is $S-$ integrable if there exist stopping times T_n such that $S^{T_n^-}$ are square-integrable semimartingales, and such that π is integrable with respect to $S^{T_n^-}$. Because the stochastic integral is defined in a way that

$$\int_0^T \pi dS^{T_n^-} \to \int_0^T \pi dS$$

almost surely, the statement of the lemma follows from the considerations before. □

References

[1] T. Moller: *On valuation and Risk Management at the interface of insurance and finance,* British Actuarial Journal, Vol. 8, No. 4 (2002)

[2] T. Moller: *Risk-minimizing hedging strategies for insurance payment processes,* Finance and Stochastics, Vol. 5, iss. 4 (2001)

[3] H. Föllmer, P. Leukert: *Quantile Hedging,* Finance & Stochastics, vol. 3, pp. 251-273 (1999)

[4] H. Föllmer, P. Leukert: *Efficient Hedging: Cost versus Shortfall Risk,* Finance & Stochastics, vol. 4, pp. 117-146 (2000)

[5] H. Witting: *Mathematische Statistik 1,* Teubner Verlag, Stuttgart (1985)

[6] H. Attouch, G. Buttazzo, G. Michaille: *Variational analysis in Sobolev and BV spaces,* MPS-SIAM Series on Optimization, Philadelphia (2005)

[7] A. Melnikov, V. Skornyakova: *Valuation and Risk Management of Equity-Linked life insurance contracts via Quantile Hedging,* Stochastic Modeling Symposium (2006)

[8] R.J. Elliott, L. Chan: *Risk measures for derivatives with Markov-modulated pure jump processes,* Asia-Pacific Finan. Markets, Vol. 13 (2006)

[9] R.J. Elliott, T.K. Siu, L. Chan: *A PDE approach for risk measures for derivatives with regime switching,* Annals of Finance, Vol. 4 (2008)

[10] P. Artzner, F. Delbaen, J. Eber, D. Heath: *Coherent measures of risk,* Mathematical Finance, Vol. 9, no. 3 (1999)

[11] Ph. Protter: *Stochastic Integration and Differential Equations,* Springer Verlag, Second edition, Heidelberg (2005)

[12] R.T. Rockafeller, S. Uryasev: *Conditional value-at risk for general loss distributions,* J. of Banking and Finance (2002)

[13] P. Forsyth, G. Labahn: *Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in Finance,* submitted to J. of Computational Finnace (2007)

[14] U. Schmock, S.E. Shreve, U. Wystup: *Valuation of Exotic Options under Shortselling Constraints,* Finance and Stochastics, Vol. 6, No. 2 (2002)