KTH Mathematics

Aspects of Cash Flow Valuation

FREDRIK ARMERIN

Doctoral Thesis
Stockholm, Sweden 2004

TRITA-MAT-04-MS-11
ISSN 1401-2286
ISRN KTH/MAT/DA--04/06--SE
ISBN 91-7283-915-5

KTH Matematik
SE-100 44 Stockholm
SWEDEN

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen fredagen den 10 dec 2004 kl 10.00 i Sal E1, Kungl Tekniska högskolan, Valhallavägen 79, Stockholm.

Tryck: Universitetsservice US AB

Abstract

This thesis consists of five papers. In the first two papers we consider a general approach to cash flow valuation, focusing on dynamic properties of the value of a stream of cash flows. The third paper discusses immunization theory, where old results are shown to hold in general deterministic models, but often fail to be true in stochastic models. In the fourth paper we comment on the connection between arbitrage opportunities and an immunized position. Finally, in the last paper we study coherent and convex measure of risk applied to portfolio optimization and insurance.

Keywords: Arbitrage, coherent and convex measures of risk, immunization, insurance, martingale methods, portfolio optimization, valuation, yield curve modelling

Acknowledgments

I thank my supervisor professor Boualem Djehiche for many interesting discussions and for his comments on earlier versions of this thesis. Professor Ingemar Kaj and assistant professor Ulf Brännlund have read parts of the manuscript and given me valuable comments.

Stockholm, November 2004
Fredrik Armerin

Contents

Chapter 1

Introduction

1.1 Summary of the thesis

This thesis deals with cash flows, and its main aim is to study how the value of a stream of cash flows behaves. As an application of cash flow valuation we consider the problem of immunization when the cash flows are of a general form. We also consider the application of coherent and convex measures of risk in portfolio optimization and insurance.

1.1.1 A general approach to cash flow valuation

The first part of the thesis consists of a general approach towards the modelling of cash flows. Consider the following simple example of cash flow valuation. At some future time T we are to be given an amount of £1. Let us denote by 0 the time of valuation. Given a constant discount rate of r and assuming continuous compounding, the value V at time 0 of getting £1 at T is given by

$$V = e^{-rT}.$$

We may now ask some questions. What will happen with V as time passes? What will happen to the value if the interest rate changes? What is the value V if we, due to the risk of bankruptcy, are not sure of getting £1 at time T? These are the types of questions we will try to answer in this thesis. Since the answers to at least the first two questions are easy to find in this simple setting, let us generalize it by considering a stream of cash flows on the fixed time interval $[0, T]$. At the deterministic times $0 \leq t_1 < t_2 < \ldots t_{n-1} < t_n \leq T$ we are given cash flows C_i, $i = 1, \ldots, n$. The value at time 0 of these cash flows is then given by

$$V = \sum_{i=1}^{n} C_i e^{-rt_i}.$$

As time passes (and hence the time of valuation) the value of the *remaining* cash flows at time $t \geq 0$ is given by

$$V(t) = \sum_{i:t_i>t} C_i e^{-r(t_i-t)} = \sum_{i=1}^{n} C_i e^{-r(t_i-t)} \mathbf{1}(t_i > t).$$

Here we have assumed the *ex dividend* principle. This means that if a payment C_i is due at time t_i, then the value at this time, $V(t_i)$, does not include C_i. The alternative would have been to assume the *cum dividend* approach, in which the payment at time t_i is included in $V(t_i)$. In this case we simply replace $\mathbf{1}(t_i > t)$ by $\mathbf{1}(t_i \geq t)$. Let us now address the first question posed above. If $t \notin \{t_1, \ldots, t_n\}$ then $dV(t) = rV(t)dt$, and if $t = t_i$ for some $i = 1, \ldots, n$ then we have $dV(t_i) = -C_i$. If we define the function $C : [0,T] \to \mathbb{R}$ by $C(t) = \sum_{i=1}^{n} C_i \mathbf{1}(t_i > t)$ then $\Delta C(t) \equiv C(t) - C(t-) = \sum_{i=1}^{n} C_i \delta_{t_i}(t)$. Hence we have

$$dV(t) = rV(t)dt - dC(t).$$

If we consider every cash flow, including the ones that have passed, in the sum we get the wealth at time t, $W(t)$, of the stream of cash flows:

$$W(t) = \sum_{i=1}^{n} C_i e^{-r(t_i-t)}.$$

In the literature this is sometimes referred to as the value of the stream of cash flows. There is no need for the payments to be lump sum payments as we have tacitly assumed so far. If we are getting payments according to the rate c, then the natural candidate for the value at time t is given by

$$V(t) = \int_t^T e^{-r(s-t)} c(s) ds.$$

We can also combine lump sum payments and payoff rates. Even more generally, we can consider modelling the cash flows by a function of finite variation. Any function f of finite variation can be used as integrator to define a Stieltjes integral. This means that we can define the value at time t of the stream of cash flows defined by the function f of finite variation as

$$V(t) = \int_{(t,T]} e^{-r(s-t)} df(s).$$

Every function of finite variation can be written as the difference of two increasing functions. Hence, we can, without loss of generality, assume f to be increasing. Since every increasing function has at most a countable number of jumps we can write $f = f^d + f^c$, where

$$f^d(t) = \sum_{0 \leq s \leq t} \Delta f(s)$$

is the jump part of f and f^c is the continuous part. Using the Lebesgue decomposition we can write $f^c = f^{ac} + f^{sc}$ where f^{ac} and f^{sc} are absolutely continuous and singular with respect to the Lebesgue measure on $[0,T]$ respectively. Letting $\dot{f}$ denote the Radon-Nikodym derivative we arrive at

$$f(t) = \sum_{0 \leq s \leq t} \Delta f(s) + \int_0^t \dot{f}(s) ds + f^{sc}(t).$$

Hence, save the singular continuous part, we gain generality by considering a mixture of lump sums and rates. Before leaving the deterministic models we remark that instead of using a constant discount rate r we could more generally consider a discount function $m(t,s)$ discounting cash flows at s to time t. We will not discuss the implications of this generalization now.

Let us return to the case when we have one payoff at time T. Instead of getting the guaranteed amount £1 we assume that we will get a random amount X at time T. We will only consider models where, if the cash flows are random, the value is given by an expectation:

$$V(T) = E\left[e^{-rT} X\right].$$

We take a pragmatic view in this thesis, not discussing why the value can be written as an expectation, but instead consider it as a postulate. If the random payment X is the money to be paid out to an insured person, in order to determine the premium an insurance company is interested in knowing this value. In this case it is natural to take expectations with respect to some probability measure estimated using available data. If on the other hand X is the payoff of some financial contract – say a call option – then given some technical assumptions the value of the call option can be written as the discounted expected value, where the expectation is taken under a so called risk-neutral probability measure. So far we have looked at the case when only the payoff is random. It may also be the case that the discount factor is random. We could assume a random interest rate, or, more generally, we could assume a random discount factor. Again letting $m(t,T)$ denote the (possibly random) discount factor by which we discount a cash flow paid at time T back to the valuation time 0 we get

$$V = E\left[m(0,T)X\right].$$

Let us now assume that the time of valuation is not 0, but some given time $t \in [0,T]$, and as above we denote by $V(t)$ the value in this case. We thus replace $m(0,T)$ by $m(t,T)$, but this is not all. When valuing X we must take into consideration that we have more information at time t than at time 0. Let us denote by E_t the expectation given all information up to and including time t.[1] Hence, we have

$$V(t) = E_t\left[m(t,T)X\right].$$

[1] This is indeed informal. The correct treatment would of course be to introduce a filtration representing the flow of information.

Given three times $s \leq u \leq t$ it can be shown that it is reasonable that m fulfills $m(s,t) = m(s,u)m(u,t)$ (compare this with the deterministic case when we have $m(s,t) = e^{-r(t-s)}$). Letting $s = 0$ gives $m(0,t) = m(0,u)m(u,t)$. If m is strictly positive we can divide by $m(0,u)$ to arrive at $m(u,t) = m(0,t)/m(0,u)$. Hence, in the case when m is strictly positive and fulfills $m(s,t) = m(s,u)m(u,t)$ it is enough to define $m(0,t)$ for every t. Let us from now on assume that this is the case. It is convenient to introduce $\Lambda_t = m(0,t)$, giving the relation

$$V(t) = E_t\left[\frac{\Lambda_T}{\Lambda_t}X\right].$$

In general the stream of cash flows we want to value does not just have one payoff. With a stream of cash flows given by times $0 \leq t_1 < \ldots < t_n \leq T$ where the random payments X_i, $i = 1,\ldots,n$, are paid, the value $V(t)$ of this stream would be

$$V(t) = E_t\left[\frac{1}{\Lambda_t}\sum_{i=1}^{n}\Lambda_{t_i}X_i\mathbf{1}(t_i > t)\right].$$

Again we have assumed the *ex dividend* principle. As in the deterministic case we can consider cash flows given by a rate, or by a process of finite variation. If the cash flows are given by the process X of finite variation the value is

$$V(t) = E_t\left[\frac{1}{\Lambda_t}\int_{(t,T]}\Lambda_s dX_s\right].$$

The stochastic integral we get when we use a process of finite variation is defined pathwisely. In the stochastic case it is also possible to define stochastic integrals in the sense of Itô. When we do this we will lose the pathwise interpretation, and we will not consider this type of models.

1.1.2 Immunization

Let us return to the valuation at time $t \in [0,T]$ of an asset giving the cash flow £1 at time T. Using a constant discount rate r the value $V(t,r)$ will be

$$V(t,r) = e^{-r(T-t)}.$$

So far we have been interested in what happens to the value as time passes. In immunization theory we want to answer the question what will happen when we change the discount factor. Both of these approaches will work well, but care has to be taken when they are combined. To see this consider a portfolio of assets with payoffs $C_i > 0$ at times $t_i \in [t,T]$. Let $K > 0$ and $T_0 \in [t,T]$ be such that

$$\sum_{i=1}^{n} C_i e^{-r(t_i-t)} = Ke^{-r(T_0-t)}$$

and

$$\sum_{i=1}^{n}(t_i - t)C_i e^{-r(t_i - t)} = (T_0 - t)Ke^{-r(T_0 - t)} \qquad (\star)$$

(such K and T_0 always exists). Define

$$\alpha_i = \frac{C_i e^{-r(t_i - t)}}{Ke^{-r(T_0 - t)}}.$$

Equation $(\star)$ can now be written

$$\sum_{i=1}^{n} \alpha_i t_i = T_0.$$

Since $\alpha_i \geq 0$ for every $i = 1, \ldots, n$ and $\sum_{i=1}^{n} \alpha_i = 1$ we can look upon $\alpha_1, \ldots, \alpha_n$ as a probability measure on $t_1, \ldots, t_n$. Define

$$g(x, y) = \sum_{i=1}^{n} C_i e^{-x(t_i - y)} - Ke^{-x(T_0 - y)}.$$

This is the value of the combined portfolio at time y and rate x. Note that $g(r, t) = 0$. For every Δr and Δt we have (using Jensen's inequality)

$$\begin{aligned}
g(r + \Delta r, t + \Delta t) &= \sum_{i=1}^{n} C_i e^{-(r+\Delta r)(t_i - t - \Delta t)} - Ke^{-(r+\Delta r)(T_0 - t - \Delta t)} \\
&= e^{(r+\Delta r)\Delta t}\left[\sum_{i=1}^{n} C_i e^{-(r+\Delta r)(t_i - t)} - Ke^{-(r+\Delta r)(T_0 - t)}\right] \\
&= Ke^{(r+\Delta r)-r(T_0 - t)}\left[\sum_{i=1}^{n} \alpha_i e^{-\Delta r(t_i - t)} - e^{-\Delta r(T_0 - t)}\right] \\
&\geq Ke^{(r+\Delta r)-r(T_0 - t)}\left[e^{-\Delta r\left(\sum_{i=1}^{n} \alpha_i t_i - t\right)} - e^{-\Delta r(T_0 - t)}\right] \\
&= 0.
\end{aligned}$$

This means that the function g has a minimum at (r, t), and that whatever change we have in r and t the new value will be non-negative. Since we did not pay anything for this position ($g(r, t) = 0$) we have an arbitrage opportunity. The reason for this is that we have used a modelling of the discount rate that allows for arbitrage. In immunization we do the same type of calculations but with the important difference that *we do not let time pass*. By fixing the time we can change the discount rate in any way we want without being afraid of considering a model which allows for arbitrage. Immunization is hence a static approach in which we can do robustness analysises of our discount rate. Looking at the previous calculations but holding t fixed shows that as long as we have a position fulfilling $g(r, t) = 0$ and Equation

($\star$) we are on the safe side regarding a possible mistake in the estimation of the discount rate. As a comparison, consider the model used by Black & Scholes to price contingent claims. In their model there is an asset with constant return r. Changing this r to, say, r' is not allowed within the model, and doing so means that we change model.

1.1.3 Risk measures

In the final part of the thesis we consider how to use convex and coherent measures of risk in portfolio optimization and insurance. Let r be the vector of returns of n financial assets having mean vector $\boldsymbol{\mu}$ and positively definite covariance matrix V. The well known Markowitz problem in portfolio theory consists of choosing the portfolio that, given a return level $\bar{r}$, minimizes the variance of the portfolio. Formally the Markowitz problem can be written

$$\left[\begin{array}{ll} \min\limits_{w\in\mathbb{R}^n} & w^T V w \\ \text{s.t.} & w^T \mathbf{1} = 1 \\ & w^T \boldsymbol{\mu} = \bar{r}. \end{array}\right.$$

It has been argued that the variance is not a good measure of risk. Instead it has been suggested that one should use a coherent or convex measure of risk. Let ρ be a coherent measure of risk having some set $\mathcal{X}$ of random variables as its domain. Given an expected return level $\bar{r}$ we now want to solve

$$\left[\begin{array}{ll} \min\limits_{w\in\mathbb{R}^n} & \rho(w^T r) \\ \text{s.t.} & w^T \mathbf{1} = 1 \\ & w^T \boldsymbol{\mu} = \bar{r}. \end{array}\right.$$

If r is normally distributed then $w^T r \stackrel{d}{=} w^T \boldsymbol{\mu} + \sqrt{w^T V w} Z$, where Z is a $N(0,1)$-distributed random variable. A coherent measure of risk ρ is always positively homogeneous and translation invariant in the sense that $\rho(X+a) = \rho(X) - a$ for every $X \in \mathcal{X}$ and $a \in \mathbb{R}$. If ρ additionally is distribution invariant ($X \stackrel{d}{=} Y \Rightarrow \rho(X) = \rho(Y)$), then

$$\rho(w^T r) = -w^T \boldsymbol{\mu} + \sqrt{w^T V w} \rho(Z).$$

One can show that if Z is a symmetric random variable then $\rho(Z) \geq 0$. If $\rho(Z) > 0$ then we see that minimizing $\rho(w^T r)$ given the return level $\bar{r}$ is equivalent to the Markowitz problem. But what if $\rho(Z) = 0$? In this case *every* vector w satisfying $w^T \mathbf{1} = 1$ and $w^T \boldsymbol{\mu} = \bar{r}$ is a solution to the problem of minimizing $\rho(w^T r)$ given $\bar{r}$. One example of a coherent risk measure having this property is the one given by $\rho(X) = E[-X]$. We study the portfolio optimization problem for coherent and convex measure of risk, both when we make distributional assumptions on r and when we do not. Finally we show that there are upper and lower bounds for every coherent and convex measure of risk. As an application of these bounds we consider an insurance company who wants to use a coherent or convex measure of risk when computing the premium, but does not know the distribution of the claims. One could then use the upper bound as a premium principle. It turns out that when the claims belong to L^2 the upper bound is a standard deviation principle, and when the claims only are assumed to belong to L^1, the upper bound is an absolute deviation principle.

1.2 Summary of the papers

1.2.1 Valuation of Cash Flows in Discrete Time

Discounted cash flow models in discrete time are considered. Under some general assumptions we show that the value of the cash flow stream can be written in three equivalent ways. We show that the discounted value tends to zero a.s. and give two cases of necessary conditions for the value process itself to converge a.s. Applications include topics from finance, economics, and life insurance.

1.2.2 Valuation of Cash Flows in Continuous Time

We model cash flows by using processes of finite variation and study the dynamic properties of the value of the stream of cash flows. The connection between the cash flow and value processes and forward-backward stochastic differential equations in a Brownian setting is discussed. Finally we consider applications of the models to real options.

1.2.3 Immunization of Deterministic and Stochastic General Cash Flows

The immunization problem when the cash flows are given by a function of finite variation and a process of finite variation respectively is studied. We show that results from classical immunization theory carry over to the case when the cash flows are modelled by a function of finite variation. This is to be contrasted with the stochastic case, where the immunization results only hold in some specific cases.

1.2.4 A Note on Immunized Portfolios, Arbitrage Opportunities and Yield Curve Modelling

In this short note we comment on the fact that it is possible to avoid problems with arbitrage opportunities in an immunized position if we look at immunization theory as a static, rather than dynamic, theory. Comparisons with the Black-Scholes model are made.

1.2.5 Coherent and Convex Measures of Risk in Portfolio Optimization and Insurance

By using coherent and convex measures of risk defined on L^p spaces we study portfolio optimization problems similar to the Markowitz problem. We show that to get a unique solution to the optimization problem additional assumptions on the risk measures is required. By using the representation theorem for coherent and convex risk measures we show that there exists upper and lower bounds on the risk measure.

Chapter 2

Cash Flow Valuation in Discrete Time

2.1 Introduction

Assume that a firm or individual is facing a stream of cash flows. These could be dividends from a stock, cash flows generated by an investment or project, or claims faced by an insurance company. What is the value today of this cash flow stream? To find the value we discount the cash flows using a suitable discount rate, take expectations and sum over time. If we call the cash flows $C_1, C_2, \dots$ and assume that the discount rate is deterministic, given by r, the discounted value at time zero is

$$V_0 = E\left[\sum_{k=1}^{\infty} \frac{C_k}{(1+r)^k}\right]$$

To make this into a dynamic model, introduce the value at time $t \geq 0$ as

$$V_t = E_t\left[\sum_{k=t+1}^{\infty} \frac{C_k}{(1+r)^{(k-t)}}\right], \tag{2.1}$$

where we let $E_t[\cdot]$ denote the expectation given information up to and including time t. By multiplying this expression with $(1+r)^{-t}$ and splitting the expectation into two parts we get

$$\frac{V_t}{(1+r)^t} = E_t\left[\sum_{k=0}^{\infty} \frac{C_k}{(1+r)^k}\right] - \sum_{k=0}^{t} \frac{C_k}{(1+r)^k}. \tag{2.2}$$

If $E\left|\sum_{k=1}^{\infty} \frac{C_k}{(1+r)^k}\right| < \infty$ then the first term on the right-hand side is a martingale and the value of the second one is known at time t. Iterating Equation (2.1) gives the relation

$$V_t = E_t\left[\frac{C_{t+1} + V_{t+1}}{1+r}\right],$$

saying that the value today is the expected discounted value of what we get tomorrow (C_{t+1}) plus the expected discounted value of having the right to the cash flow stream $C_{t+2}, C_{t+3}, \ldots$ (which is the definition of V_{t+1}). By continued iterations we get for any $T > t$

$$V_t = E_t\left[\sum_{k=t+1}^{T} \frac{C_k}{(1+r)^{k-t}}\right] + E_t\left[\frac{V_T}{(1+r)^T}\right].$$

If we impose the condition that the last term in the right-hand side of the previous equation goes to zero as T goes to infinity we are back to Equation (2.1) We see from Equation (2.2) that if we let t go to infinity, then the discounted value $V_t/(1+r)^t$ tends to 0 a.s. A subsequent question is now what will happen to the value V_t when we let time go to infinity. It turns out that this convergence depends on the behavior of both the discount factors and the cash flows. The idea of rewriting the value equation (2.1) as to identify the martingale embedded within comes from life insurance. There the expected discounted value of the cash flows is known as the retroperspective reserve. The fact that we can decompose the discounted value as the difference of a martingale and an adapted process give us a way to prove Hattendorff's Theorem. Recently valuation using real options has gained increasing interest. In these models either the value or the underlying cash flow is modelled as a stochastic process. In the latter case the question of how the dynamical properties of the cash flows influence the dynamics of the value process is important. For references and more concrete examples see Section 2.3.1. The purpose of this paper is to prove the results indicated above in a more general setting. In Section 2 we define the cash flow process as any a.s. finite adapted process and the discount process as an adapted process, fulfilling a consistency relation. While we do not comment much upon the cash flows, the discount process and its equivalent forms, is discussed in some detail. In Section 3 we define the value process. We discuss some properties of it and then state and prove that there are three equivalent forms in which we can express the value process. We then turn to the problem of the convergence of the value process. Although the discounted value tends to 0, the convergence of the value itself depends on both the the cash flows and the discount rates. We give two propositions containing necessary condition for the convergence of the value process. The last part of Section 3 contains the case when the cash flows and/or the value process is evaluated at a stopping time. We find that the earlier result easily extends to this situation.

2.2 General definitions

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}}, P)$ be a complete filtered probability space. We will assume that $\mathcal{F}_0$ is the trivial σ-algebra augmented with all null sets of $\mathcal{F}$ and that $\mathcal{F}_\infty = \mathcal{F}$, where $\mathcal{F}_\infty = \bigvee_{t\geq 0} \mathcal{F}_t$. We will use the conventions $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_{++} = (0, \infty)$.

Definition 2.1 *A* cash flow process $(C_t)_{t\in\mathbb{N}}$ *is a process adapted to the filtration* $(\mathcal{F}_t)$ *and such that for each* $t \in \mathbb{N}$, $|C_t| < \infty$ *a.s. A cash flow process that is non-negative a.s. will be referred to as a* dividend process.

The discount process tells us how to discount future payments.

Definition 2.2 *A* discount process *is a process* $m : \mathbb{N} \times \mathbb{N} \times \Omega \to \mathbb{R}$ *satisfying*

(i) $0 < m(s,t) < \infty$ *a.s. for every* $s, t \in \mathbb{N}$,

(ii) $m(s,t,\omega)$ *is* $\mathcal{F}_{\max(s,t)}$*-measurable for every* $s,t \in \mathbb{N}$, *and*

(iii) $m(s,t) = m(s,u)m(u,t)$ *a.s for every* $s,u,t \in \mathbb{N}$.

A discount process fulfilling

(i') $0 < m(s,t) \leq 1$ *a.s. for every* $s,t \in \mathbb{N}$ *with* $s \leq t$

will be referred to as a normal *discount process.*

As a short hand notation we will write $m_t \equiv m(0,t)$, $t \in \mathbb{N}$. Implied by the assumptions on the discount factors is the fact that $m(t,t) = 1$. This is seen by letting $s = u = t$ in (iii) together with the fact that $m > 0$ a.s. Now let $s < t$. We interpret $m(s,t)$ as the (stochastic) value at time s of getting one unit of currency at t, and analogously we interpret $m(t,s)$ as the growth of one unit of currency, invested at time t, at time s. In this latter case we should rather call m an accumulation factor. That we allow $m(s,t)$ with $s > t$ is because we want to incorporate insurance models into our framework. In life insurance applications we need to be able to both discount and accumulate cash flows. Condition (iii) in the definition could be seen as a consistency or no arbitrage condition, see Norberg [10]. The following two examples of discount factors are 'typical' (see Lemma 2.5 below).

Example 2.3 *Let* $r \in \mathbb{R}$. *Then it is easy to verify that*

$$m(s,t) = e^{-r(t-s)}$$

is a (deterministic) discount process. It is not difficult to see that m *is a normal discount process if and only if* $r \geq 0$. □

We can generalize this example to get a stochastic discount process.

Example 2.4 *Let*

$$m(s,t) = \exp\left(-\sum_{k=s+1}^{t} f_k\right),$$

with $(f_k)_{k\in\mathbb{N}}$ *an adapted process that is finite a.s. As in the previous example it is immediate that* m *fulfills the requirements of a discount process. The requirement* $f_k \geq 0$ *a.s. will make* m *a normal discount process.* □

Assumption (iii) in the definition of the discount process gives a lot of structure, as is seen in the following lemma.

Lemma 2.5 *m is a discount process if and only if it can be written*

$$m(s,t) = \frac{\Lambda(t)}{\Lambda(s)}, \quad \textit{a.s. for all } s,t \in \mathbb{N}, \tag{2.3}$$

where $(\Lambda(t))_{t\in\mathbb{N}}$ *is an a.s. strictly positive and finite adapted process*

Proof. We begin with the 'if' part. Obviously

$$m(s,u)m(u,t) = \frac{\Lambda(u)}{\Lambda(s)} \cdot \frac{\Lambda(t)}{\Lambda(u)} = \frac{\Lambda(t)}{\Lambda(s)} = m(s,t) \text{ a.s. for all } s,u,t \in \mathbb{N}.$$

The fact that $\Lambda(t) > 0$ a.s. implies that $m(s,t) > 0$ a.s. Since $\Lambda(t)$ is $\mathcal{F}_t$-measurable for every $t \in \mathbb{N}$, $m(s,t)$ will be $\mathcal{F}_{\max(s,t)}$-measurable for all s,t in $\mathbb{N}$. For the 'only if' part we begin by noting that since $m(0,t) > 0$ a.s. for every $t \in \mathbb{N}$ we have

$$m(0,t) = m(0,s)m(s,t) \text{ a.s. if and only if } m(s,t) = \frac{m(0,t)}{m(0,s)} \text{ a.s.}$$

Now let $\Lambda(t) := m(0,t)$. It is easily seen that this choice of $\Lambda(t)$ fulfills the desired requirements. □

The connection to Example 2.4 above becomes more transparent if we write (2.3) as

$$m(s,t) = \exp\left(-(\ln \Lambda_s - \ln \Lambda_t)\right) = \exp\left(-\sum_{k=s}^{t-1} \ln \frac{\Lambda_k}{\Lambda_{k+1}}\right).$$

The process Λ is known as a deflator. If it is the price of a traded asset, it is called a numeraire. In the theory of arbitrage pricing one can show that the existence of a discount process is equivalent to a condition ruling out arbitrage strategies. The exact condition is that the stock price process should satisfy the condition of 'no free lunch with bounded risk' (NFLBR). Intuitively this means that there is no possibility of having strategies such that the profit of the strategy can be arbitrarily large while the maximum loss of using the strategy is restricted to 1 monetary unit. The definition of (NFLBR) and its connection to pricing is discussed in Schachermayer [14]. See also Section 4.C in Duffie [5] and Chapter 7 in Pliska [13] for no arbitrage pricing with an infinite discrete time horizon. Since the cash flows generated by a project or the claims in life insurance typically are not traded, conditions for the existence of a martingale measure are not a relevant question for us.

Definition 2.6 *The* discount rate *or the* instantaneous rate *at time t implied by the discount process, denoted* r_t *for* $t = 1, 2, \ldots$, *is defined as*

$$r_t = \frac{1}{m(t-1,t)} - 1 = \frac{m_{t-1}}{m_t} - 1 = \frac{\Lambda(t-1)}{\Lambda(t)} - 1, \; t = 1, 2, \ldots$$

where Λ *is the deflator associated with m.*

The advantage of using the instantaneous rates, which uniquely determines the discount process, is that a requirement on the rates is often more easy to interpret economically than a requirement put on the discount process. The following lemma contains some facts relating the rate process and discount process.

Lemma 2.7 *Let m be a discount process and let r be the discount rate implied by m. Then the following holds:*

(i) $-1 < r_t < \infty$, $t \in \mathbb{N}$

(ii) $r \geq 0$ *if and only if* m *is a normal discount process.*

(iii) For any given $\lambda > 0$ *we have for* $t \in \mathbb{N}$

$$0 < \lambda \leq r_t \quad \textit{if and only if} \quad 0 < m_t \leq e^{-t\ln(1+\lambda)}.$$

(iv) The instantaneous rate process and the discount process uniquely determine each other.

Proof. Facts (i) and (ii) are immediate from the definition. To get (iii) we have the following implications for any $\lambda > 0$ and $t \in \mathbb{N}$:

$$\lambda \leq r_t \quad \Rightarrow \quad \lambda \leq \frac{m_{t-1}}{m_t} - 1 \quad \Rightarrow \quad m_t \leq \frac{1}{1+\lambda} m_{t-1}.$$

Iterating this gives

$$m_t \leq \left(\frac{1}{1+\lambda}\right)^t = e^{-t\ln(1+\lambda)}.$$

To go in the other direction we see that using the definition of r together with the fact that $\lambda > 0$ gives the desired result. For (iv) finally we see that given m the discount rate process r is determined uniquely. The opposite conclusion is clear from the following:

$$m(t,k) = \prod_{\ell=t+1}^{k} m(\ell-1,\ell) = \prod_{\ell=t+1}^{k} \frac{1}{1+r_\ell}.$$

□

2.3 Valuation

Definition 2.8 *Given a cash flow process* $(C_t)_{t\in\mathbb{N}}$ *and a discount process* $(m(s,t) : s,t \in \mathbb{N})$ *we define for* $t \in \mathbb{N}$ *the* value process *as*

$$V_t = E\left[\sum_{k=t+1}^{\infty} C_k m(t,k) \,\middle|\, \mathcal{F}_t\right].$$

The value process is defined *ex dividend*, meaning that we include cash flows from time $t+1$ and onwards in the value at time t. It would be possible to define it *cum dividend*, thus also including the cash flow at time t, but since the *ex dividend* version is the most usual in financial texts we prefer it. See e.g. Campbell et al [3] or Cuthbertson [4] for more details on this issue.

Recall that the only conditions we have put on the cash flow process is that $|C_t| < \infty$ a.s. for $t \in \mathbb{N}$. A natural question to ask is when the value process will be finite a.s. The following lemma offers a sufficient condition for this.

Lemma 2.9 *If C is a cash flow process and $E\left[\left|\sum_{k=1}^{\infty} C_k m_k\right|\right] < \infty$ a.s. then $|V_t| < \infty$ a.s. for all $t \in \mathbb{N}$.*

Proof. Since $E\left[\left|\sum_{k=1}^{\infty} C_k m_k\right|\right] < \infty$, the following conditional expectations are well defined: For $t \in \mathbb{N}$

$$\begin{aligned} |V_t| = \left|E\left[\sum_{k=t+1}^{\infty} C_k m(t,k)\middle|\mathcal{F}_t\right]\right| &\leq \frac{1}{m(0,t)} E\left[\left|\sum_{k=1}^{\infty} C_k m(0,k) - \sum_{k=1}^{t} C_k m(0,k)\right|\middle|\mathcal{F}_t\right] \\ &\leq \frac{1}{m_t}\left(E\left[\left|\sum_{k=1}^{\infty} C_k m_k\right|\middle|\mathcal{F}_t\right] + \left|\sum_{k=1}^{t} C_k m_k\right|\right) \\ &< \infty \text{ a.s.} \end{aligned}$$

□

We immediately get the following corollary for a dividend process.

Corollary 2.10 *If C is a dividend process such that $V_0 < \infty$, then $V_t < \infty$ a.s. for every $t \in \mathbb{N}$.*

Proof. Since $C_t \geq 0$ a.s. for every $t \in \mathbb{N}$ when C is a dividend process and $\mathcal{F}_0$ is the trivial σ-algebra augmented with the null sets

$$E\left[\left|\sum_{k=1}^{\infty} C_k m_k\right|\right] = E\left[\sum_{k=1}^{\infty} C_k m_k\right] = V_0 < \infty,$$

and the previous lemma applies. □

We now proceed by rewriting the value process. Note that since m_t is $\mathcal{F}_t$-measurable for all $t \in \mathbb{N}$ we have

$$V_t = E\left[\sum_{k=t+1}^{\infty} C_k m(t,k)\middle|\mathcal{F}_t\right] = \frac{1}{m_t} E\left[\sum_{k=t+1}^{\infty} C_k m_k\middle|\mathcal{F}_t\right].$$

By multiplying the expression for V_t by m_t we get

$$V_t m_t = E\left[\sum_{k=0}^{\infty} C_k m_k \middle| \mathcal{F}_t\right] - \sum_{k=0}^{t} C_k m_k.$$

Note that $V_t m_t$ is the value at time t discounted back to time 0. It is well known that if X is a random variable with $E|X| < \infty$ then $E[X|\mathcal{F}_t]$, $t = 1, 2, \ldots$ is a uniformly integrable (UI) martingale. Thus, if $E[|\sum_{k=0}^{\infty} C_k m_k|] < \infty$ then $E[\sum_{k=0}^{\infty} C_k m_k|\mathcal{F}_t]$ is a UI martingale. This and other facts characterizing the discounted value process are summarized in the following proposition.

Proposition 2.11 *Let C and m be a cash flow and discount process respectively. If $E|\sum_{k=1}^{\infty} C_k m_k| < \infty$ then the discounted value process $(V_t m_t)$ can be written*

$$V_t m_t = M_t - A_t, \ t \in \mathbb{N},$$

where M is a UI martingale and A is an adapted process. Furthermore

$$\lim_{t\to\infty} V_t m_t = 0 \ a.s.$$

Proof. We notice that $|\sum_{k=0}^{\infty} C_k m_k| < \infty$ a.s. since $E[|\sum_{k=0}^{\infty} C_k m_k|] < \infty$. Now let

$$\begin{aligned} M_t &= E\left[\sum_{k=1}^{\infty} C_k m_k \middle| \mathcal{F}_t\right], \ t \in \mathbb{N}, \\ A_t &= \sum_{k=1}^{t} C_k m_k, \ t \in \mathbb{N}. \end{aligned}$$

It is then immediate that $V_t m_t = M_t - A_t$. Since $E[|\sum_{k=1}^{\infty} C_k m_k|] < \infty$ M is a UI martingale and we see that A is adapted. We know that (Williams [15] p. 134) the UI martingale will converge to $E[\sum_{k=0}^{\infty} C_k m_k|\mathcal{F}_\infty] = \sum_{k=0}^{\infty} C_k m_k$ a.s. as $t \to \infty$. This yields

$$\lim_{t\to\infty} V_t m_t = \lim_{t\to\infty} M_t - \lim_{t\to\infty} A_t = 0,$$

since $M_\infty = A_\infty = \sum_{k=0}^{\infty} C_k m_k$ is finite a.s. □

The following theorem characterizes the relation between C, m and V in terms of their values and differences, giving three equivalent forms of defining the value process.

Theorem 2.12 *Let C be a cash flow process and m a discount process such that $E[\sum_{k=1}^{\infty} |C_k| m_k] < \infty$. Then the following three statements are equivalent.*

(i) For every $t \in \mathbb{N}$

$$V_t = E\left[\sum_{k=t+1}^{\infty} C_k m(t,k) \middle| \mathcal{F}_t\right].$$

(ii) *(a)* *For every* $t \in \mathbb{N}$

$$M_t = V_t m_t + \sum_{k=1}^{t} C_k m_k$$

is a UI martingale, and

(b) $V_t m_t \to 0$ *a.s. when* $t \to \infty$.

(iii) *For every* $t \in \mathbb{N}$

(a) $V_t = E\left[m(t,t+1)(C_{t+1} + V_{t+1})|\mathcal{F}_t\right]$, *and*

(b) $\lim_{T\to\infty} E\left[m(t,T)V_T|\mathcal{F}_t\right] = 0$.

Proof. First of all we note that $E\left|\sum_{k=1}^{\infty} C_k m_k\right| \leq E\left[\sum_{k=1}^{\infty} |C_k| m_k\right] < \infty$, so $\left|\sum_{k=1}^{\infty} C_k m_k\right| < \infty$ a.s. We will show (i) ⇔ (ii) and (i) ⇔ (iii)

(i) ⇔ (ii): The 'only if' part follows from Proposition 2.11. For the 'if' part write the expression in (ii) (a) as $-m_{k+1}C_{k+1} = m_{k+1}V_{k+1} - m_k V_k - M_{k+1} + M_k$ and sum from t to $T-1$:

$$-\sum_{k=t+1}^{T} m_k C_k = m_T V_T - m_t V_t - M_T + M_t.$$

Letting $T \to \infty$ the term $m_T V_T \to 0$ a.s. by the assumption and $M_T \to M_\infty$ a.s. from the convergence result of UI martingales (Williams [15] p. 134). Thus we have

$$V_t m_t = \sum_{k=t+1}^{\infty} C_k m_k - M_\infty + M_t \text{ a.s.}$$

The convergence result concerning UI martingales also ensures the relation $E\left[M_\infty|\mathcal{F}_t\right] = M_t$ a.s. Taking conditional expectations with respect to $\mathcal{F}_t$ and using the definition of discount factors yields

$$V_t = E\left[\sum_{k=t+1}^{\infty} C_k m(t,k)\middle|\mathcal{F}_t\right].$$

(i) ⇔ (iii): We begin with the 'only if' part. Fix a $t \in \mathbb{N}$. We get

$$\begin{aligned} V_t &= E\left[m(t,t+1)C_{t+1} + \sum_{k=t+2}^{\infty} C_k m(t,k)\middle|\mathcal{F}_t\right] \\ &= E\left[m(t,t+1)C_t + m(t,t+1)\sum_{k=t+2}^{\infty} C_k m(t+1,k)\middle|\mathcal{F}_t\right] \\ &= E\left[m(t,t+1)(C_{t+1} + V_{t+1})|\mathcal{F}_t\right]. \end{aligned}$$

Now let $T \geq t$. From $V_T = E\left[\sum_{k=T+1}^{\infty} C_k m(T,k) | \mathfrak{F}_T\right]$ we get

$$E\left[m(t,T)V_T|\mathfrak{F}_t\right] = E\left[\sum_{k=T+1}^{\infty} C_k m(t,k)\middle|\mathfrak{F}_t\right] = \frac{1}{m_t} E\left[\sum_{k=T+1}^{\infty} C_k m_k\middle|\mathfrak{F}_t\right].$$

Since

$$\left|\sum_{k=T+1}^{\infty} C_k m_k\right| \leq \sum_{k=1}^{\infty} |C_k| m_k$$

and the last random variable is integrable by assumption we get, for every $t \in \mathbb{N}$ and $A \in \mathfrak{F}_t$,

$$\lim_{T\to\infty} E\left[m(t,T)V_T \mathbf{1}_A\right] = E\left[\lim_{T\to\infty} m(t,T)V_T\mathbf{1}_A\right] = 0.$$

To prove the 'if' part we iterate (iii) (a) to get

$$\begin{aligned} V_t &= E\left[\sum_{k=t+1}^{T} C_k m(t,k) + m(t,T)V_T\middle|\mathfrak{F}_t\right] \\ &= \frac{1}{m_t} E\left[\sum_{k=t+1}^{T} C_k m_k\middle|\mathfrak{F}_t\right] + E\left[m(t,T)V_T|\mathfrak{F}_t\right]. \end{aligned}$$

When we let $T \to \infty$ the the last term tend to 0 a.s. from (iii) (b). Since

$$\left|\sum_{k=t+1}^{T} C_k m_k\right| \leq \sum_{k=1}^{\infty} |C_k| m_k$$

and the last random variable is integrable by assumption we get

$$V_t = E\left[\sum_{k=t+1}^{\infty} C_k m(t,k)\middle|\mathfrak{F}_t\right]$$

by using the Theorem of Dominated Convergence. □

Before proceeding we make some comments on the previous theorem. We have written conditions (ii) (a) and (iii) (b) on the form in the theorem because of its convenient form. A more intuitive way, from an economical/financial point of view, is to write condition (ii) (a) as

$$\Delta V_t = r_t V_{t-1} - C_t + \frac{1}{m_t}\Delta M_t,$$

where $r_t = m_{t-1}/m_t - 1$ is the instantaneous rate, and condition (iii) (a) as

$$V_t = E\left[\frac{V_{t+1} + C_{t+1}}{1 + r_{t+1}}\middle|\mathfrak{F}_t\right].$$

Note that if m_t is predictable, then $(1/m_t)\Delta M_t$ is a martingale difference, and we have $E\left[\Delta V_t|\mathfrak{F}_{t-1}\right] = r_t V_{t-1} - E\left[C_t|\mathfrak{F}_{t-1}\right]$.

2.3.1 Examples

We will now discuss some well known relations from finance, economics and insurance where the use of Theorem 2.12 is needed. In these applications often some assumptions on the cash flows and/or the discount processes are made. Theorem 2.12 however shows that the reasoning can be made under quite mild assumptions.

It is a well-known fact from arbitrage pricing that the discounted gains process should be a martingale under an equivalent martingale measure In our setting the UI martingale M represents the discounted gains process. See Duffie [5] and Pliska [13] for theory and applications of no arbitrage pricing in discrete time.

If we define

$$L_t = M_t - M_{t-1} = V_t m_t - V_{t-1} m_{t-1} + C_t m_t,$$

then L will be a sequence of martingale differences and we will especially have $E[L_t L_s] = 0$ for all $s, t \in \mathbb{N}$. If the cash flows are interpreted as losses faced by an insurance company, then L_t is the discounted annual loss in the time period $(t-1, t]$. The fact that the discounted annual losses are uncorrelated is in life insurance known as Hattendorff's Theorem. See Papatriandafylou & Waters [12] for this result and more on the same theme. Early examples of these methods in life insurance are Bühlmann [1] and Gerber [9]. We remark that in life insurance applications the value process V is known as the prospective reserve. The value Q_t at time t of a cash flow stream C of insurance claims is then defined to be

$$Q_t = \frac{1}{m_t} E\left[\sum_{k=1}^{\infty} C_k m_k \middle| \mathcal{F}_t\right] = \sum_{k=1}^{t} m(t,k) C_k + E\left[\sum_{k=t+1}^{\infty} C_k m_k \middle| \mathcal{F}_t\right] =: A_t + R_t,$$

where A_t is the accumulated payments and R_t is the prospective reserve (i.e. what we call the value process). See Bühlmann [2] and Bühlmann's contribution in [11].

In financial economics and econometric models, the starting point is often (iii) (a) in Theorem 2.12. The return of a stock from time t to time $t+1$ is defined as

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} - 1,$$

where P_t and P_{t+1} is the price of the stock at time t and $t+1$ respectively and D_{t+1} is the dividend per share at $t+1$. Taking the conditional expectation with respect to $\mathcal{F}_t$ gives

$$P_t = E\left[\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}} \middle| \mathcal{F}_t\right]; \tag{2.4}$$

which is (iii) (a) with renamed processes. By iterating this we get

$$P_t = E\left[\sum_{k=t+1}^{T} D_k \prod_{\ell=t+1}^{k} \left(\frac{1}{1+R_\ell}\right) \middle| \mathcal{F}_t\right] + E\left[P_T \prod_{\ell=t+1}^{T} \left(\frac{1}{1+R_\ell}\right) \middle| \mathcal{F}_t\right].$$

To be able to write the stock price at time t as the discounted sum of all future dividends the second term in the equation above has to go to zero a.s. This condition, (iii) (b) in Theorem 2.12, is known as the transversality condition. Now let us look for solutions to Equation (2.4), dropping all other assumptions on the behavior of the solution. In this case there is no longer a unique solution. We call the solution with the transversality condition imposed P^D. Obviously this will be a solution even when we look for solutions only to (2.4). Now we have the following fact: Any solution P to Equation (2.4) can be written

$$P_t = P_t^D + \frac{Z_t}{m_t}, \ t \in \mathbb{N},$$

where Z has the martingale property and $m_t = \prod_{k=1}^{t} \left(\frac{1}{1+R_k}\right)$. To see this, let P be any solution to (2.4). Then

$$P_t - P_t^D = E\left[\frac{P_t - P_{t+1}^D}{1 + R_{t+1}} \middle| \mathcal{F}_t\right]$$

if and only if

$$(P_t - P_t^D)m_t = E\left[\left(P_{t+1} - P_{t+1}^D\right) m_{t+1} \middle| \mathcal{F}_t\right],$$

implying that $(P - P^D)m$ has the martingale property. The solution P^D is known as the fundamental value or the bubble free solution and Z/m is called a rational bubble. The process Z/m is called a bubble since its presence yields prices that are higher than the fundamental value, and it is 'rational' in the sense that it is not inconsistent with rational expectations. Campbell et al [3] and Cuthbertson [4] discuss rational bubbles from both a theoretical and empirical point of view.

Finally we mention the important subclass of Markov models. By assuming an underlying Markov process driving the cash flows and the discount rate the general formula for the value process can be further simplified. Much of this can be found and is commented on in Duffie [6]. There the close connection between Markov pricing and semigroups is pointed out. For the semigroup approach see Garman [8] and references therein. See also the general texts Duffie [5] and Pliska [13].

2.3.2 Asymptotic behavior of the value process

We know that $V_t m_t \to 0$ a.s., but what will happen to V_t when $t \to \infty$? We will present two results showing that V_t can, given some conditions, converge to 'almost anything' in ways which will be specified below. The essential assumption is that we have a strong law of large numbers for the sequence $\log(1 + r_t)$. Roughly this means that the discount process behaves like $e^{-\lambda t}$ for some $\lambda > 0$ when t is large.

Proposition 2.13 *Let C and m be a cash flow process and discount process respectively. If* *(i) There exists a constant* $\lambda > 0$ *such that*

$$\frac{m_t}{e^{-\lambda t}} \to 1 \text{ a.s. as } t \to \infty,$$

(ii) $C_t \to C_\infty \in L^1$ *a.s. as* $t \to \infty$*, and*

(iii) there exists an integrable random variable Z *s.t. for all* $t \geq 0$

$$\left| C_t \frac{m_t}{e^{-\lambda t}} \right| \leq Z \text{ a.s.}$$

then

$$V_t \to \frac{e^{-\lambda}}{1 - e^{-\lambda}} C_\infty \text{ a.s. as } t \to \infty$$

Proof. We have $m_t \to e^{-\lambda t}$ a.s. as $t \to \infty$.

$$\begin{aligned} V_t &= E\left[\sum_{k=t+1}^{\infty} C_k m(t,k) \middle| \mathfrak{F}_t \right] = \frac{1}{m_t} E\left[\sum_{k=t+1}^{\infty} C_k m_k \middle| \mathfrak{F}_t \right] \\ &= \frac{e^{-\lambda t}}{m_t} E\left[\sum_{k=1}^{\infty} e^{-\lambda k} C_{k+t} \frac{m_{k+t}}{e^{-\lambda(k+t)}} \middle| \mathfrak{F}_t \right] \end{aligned} \tag{2.5}$$

Now from (iii) above

$$\left| \sum_{k=1}^{\infty} e^{-\lambda k} C_{k+t} \frac{m_{k+t}}{e^{-\lambda(k+t)}} \right| \leq \sum_{k=1}^{\infty} e^{-\lambda k} \left| C_{k+t} \frac{m_{k+t}}{e^{-\lambda(k+t)}} \right| \leq Z \sum_{k=1}^{\infty} e^{-\lambda k} = Z \frac{e^{-\lambda}}{1 - e^{-\lambda}},$$

implying that

$$E\left[\left| \sum_{k=1}^{\infty} e^{-\lambda k} C_{k+t} \frac{m_{k+t}}{e^{-\lambda(k+t)}} \right| \right] \leq E[Z] \frac{e^{-\lambda}}{1 - e^{-\lambda}} < \infty$$

We now use the Dominated Convergence Theorem for conditional expectations (see e.g. Durrett [7] p. 264). To do this, first note that

$$\lim_{t \to \infty} \sum_{k=1}^{\infty} e^{-\lambda k} C_{k+t} \frac{m_{k+t}}{e^{-\lambda(k+t)}} = \sum_{k=1}^{\infty} e^{-\lambda k} C_\infty = \frac{e^{-\lambda}}{1 - e^{-\lambda}} C_\infty,$$

where we have used the Dominated Convergence Theorem. Now it follows from the Theorem of Dominated Convergence for conditional expectations that when $t \to \infty$,

$$E\left[\sum_{k=1}^{\infty} e^{-\lambda k} C_{k+t} \frac{m_{k+t}}{e^{-\lambda(k+t)}} \middle| \mathfrak{F}_t \right] \to E\left[\frac{e^{-\lambda}}{1 - e^{-\lambda}} C_\infty \middle| \mathfrak{F}_\infty \right] = \frac{e^{-\lambda}}{1 - e^{-\lambda}} C_\infty \text{ a.s.}$$

Now let $t \to \infty$ in Equation (2.5). Since $\frac{e^{-\lambda t}}{m_t} \to 1$ a.s. it follows that $V_t \to \frac{e^{-\lambda}}{1-e^{-\lambda}} C_\infty$ a.s. when $t \to \infty$, and the proposition is proved. □

Corollary 2.14 *Let C and m be a dividend and discount process respectively. If $0 \leq C_t \uparrow C_\infty$ a.s. as $t \to \infty$ with C_∞ being integrable, and there exists a constant $\lambda > 0$ such that*

$$1 \geq \frac{m_t}{e^{-\lambda t}} \to 1 \text{ a.s. as } t \to \infty,$$

then $V_t \to C_\infty$ a.s. as $t \to \infty$.

Proof. We have

$$\left| C_t \frac{m_t}{e^{-\lambda t}} \right| = C_t \frac{m_t}{e^{-\lambda t}} \leq C_\infty,$$

and since C_∞ is integrable the previous proposition applies. □

Proposition 2.13 has the unsatisfactory integrability condition (iii). The following result does not need this, but is on the other hand another kind of result. It says that given an integrable random variable X, there exists a cash flow process such that associated value processes converges to X a.s.

Proposition 2.15 *Let m be a discount process, and let X be an integrable random variable. If there exists a constant $\lambda > 0$ such that*

$$\frac{m_t}{e^{-\lambda t}} \to 1 \text{ a.s. as } t \to \infty$$

then there exists a cash flow process such that

$$V_t \to X \quad a.s.$$

Proof. Take $\lambda > 0$ such that $m_t/e^{-\lambda t} \to 1$ a.s. as $t \to \infty$, and fix $t \geq 0$. For $k \geq t$ let

$$C_k = \frac{E\left[X|\mathcal{F}_k\right] e^{-\lambda k}(1 - e^{-\lambda})}{m_k e^{-\lambda}}.$$

Now,

$$\begin{aligned} V_t &= E\left[\sum_{k=t+1}^{\infty} \frac{E\left[X|\mathcal{F}_k\right] e^{-\lambda k}(1 - e^{-\lambda})}{m_k e^{-\lambda}} m(t,k) \middle| \mathcal{F}_t\right] \\ &= \frac{1 - e^{-\lambda}}{m_t e^{-\lambda}} E\left[\sum_{k=1}^{\infty} E\left[X|\mathcal{F}_{t+k}\right] e^{-\lambda(t+k)} \middle| \mathcal{F}_t\right] \\ &= \frac{1 - e^{-\lambda}}{m_t e^{-\lambda}} E\left[X|\mathcal{F}_t\right] \sum_{k=1}^{\infty} e^{-\lambda(t+k)} \\ &= \frac{e^{-\lambda t}}{m_t} E\left[X|\mathcal{F}_t\right] \to X \text{ a.s.} \end{aligned}$$

as $t \to \infty$ since $e^{-\lambda t}/m_t \to 1$ a.s. and $E[X|\mathcal{F}_t] \to E[X|\mathcal{F}_\infty] = X$ a.s. when $t \to \infty$. The interchange of summation and conditional expectation is justified by the Fubini theorem. To see this first note that for $A \in \mathcal{F}_t$

$$\begin{aligned} E\left|\mathbf{1}_A E[X|\mathcal{F}_{t+k}]\right| &= E\left[E\left[\mathbf{1}_A |E[X|\mathcal{F}_{t+k}]|\,|\mathcal{F}_t\right]\right] \\ &= E\left[\mathbf{1}_A E\left[|E[X|\mathcal{F}_{t+k}]|\,|\mathcal{F}_t\right]\right] \\ &\leq E\left[\mathbf{1}_A E\left[E\left[|X|\,|\mathcal{F}_{t+k}\right]|\mathcal{F}_t\right]\right] \\ &= E\left[\mathbf{1}_A E\left[|X|\,|\mathcal{F}_t\right]\right]. \end{aligned}$$

Thus for any $A \in \mathcal{F}_t$ we get

$$\begin{aligned} E\left[\sum_{k=0}^{\infty}\left|\mathbf{1}_A E[X|\mathcal{F}_{t+k}]\, e^{-\lambda(t+k)}\right|\right] &= \sum_{k=0}^{\infty} E\left|\mathbf{1}_A E[X|\mathcal{F}_{t+k}]\right| e^{-\lambda(t+k)} \\ &\leq \sum_{k=0}^{\infty} e^{-\lambda(t+k)} E\left[\mathbf{1}_A E\left[|X|\,|\mathcal{F}_t\right]\right] \\ &= e^{-\lambda t} E\left[\mathbf{1}_A E\left[|X|\,|\mathcal{F}_t\right]\right] \sum_{k=0}^{\infty} e^{-\lambda k} \\ &= \frac{e^{-\lambda t}}{1-e^{-\lambda}} E\left[\mathbf{1}_A E\left[|X|\,|\mathcal{F}_t\right]\right] < \infty, \end{aligned}$$

which justifies the interchange of expectation and summation. □

2.3.3 Stopping the cash flow and value process

Theorem 2.12 on the three equivalent representations of the value process concerns the value at deterministic times. It also assumes that the cash flow stream is defined for all $t \geq 0$. In some cases we would like to consider the value at a stopping time and/or the cash flow process stopped at some stopping time. By utilizing the fact that the martingale $M_t = V_t m_t + \sum_{k=1}^{t} C_k m_k$ from Theorem 2.12 is uniformly integrable we can get the following result.

Proposition 2.16 *Let C be a cash flow process and let m be a discount process such that $E\left[\sum_{k=t+1}^{\infty} |C_k| m(t,k)\right] < \infty$ for every $t \in \mathbb{N}$. Further let τ and σ be $(\mathcal{F}_t)$-stopping times such that $\sigma \leq \tau$ a.s. Then the following two statements are equivalent*

(i) We have

$$V_\sigma = E\left[\sum_{k=\sigma+1}^{\tau} C_k m(\sigma,k) + V_\tau m(\sigma,\tau)\mathbf{1}_{\tau<\infty}\middle|\mathcal{F}_\sigma\right] \quad \text{on } \{\sigma < \infty\}.$$

(ii) *(a)* *For every* $t \in \mathbb{N}$

$$M_t = V_t m_t + \sum_{k=1}^{t} C_k m_k$$

is a UI martingale, and

(b) $V_t m_t \to 0$ *a.s. when* $t \to \infty$.

Proof. We begin with the implication (ii) $\Rightarrow$ (i). The stopping time τ may be unbounded so we consider the stopping times $\tau \wedge n$, where $n \in \mathbb{N}$. We get

$$M_{\tau\wedge n} = V_{\tau\wedge n} m_{\tau\wedge n} + \sum_{k=1}^{\tau\wedge n} C_k m_k. \tag{2.6}$$

Now,

$$V_{\tau\wedge n} m_{\tau\wedge n} \xrightarrow{\text{a.s.}} V_\tau m_\tau \mathbf{1}_{\tau<\infty},$$

as $n \to \infty$ since $V_n m_n \mathbf{1}_{\tau=\infty} \to 0$ a.s. By letting $n \to \infty$ in Equation (2.6) we thus get

$$M_\tau = \sum_{k=1}^{\tau} C_k m_k + V_\tau m_\tau \mathbf{1}_{\tau<\infty}.$$

Since M is uniformly integrable we can take the conditional expectation of M_τ with respect to the σ-algebra $\mathcal{F}_\sigma$ to get on $\{\sigma < \infty\}$

$$\begin{aligned}
\sum_{k=1}^{\sigma} C_k m_k + V_\sigma m_\sigma &= M_\sigma = E\left[M_\tau | \mathcal{F}_\sigma\right] \\
&= E\left[\sum_{k=1}^{\tau} C_k m_k + V_\tau m_\tau \mathbf{1}_{\tau<\infty} \middle| \mathcal{F}_\sigma\right] \\
&= \sum_{k=1}^{\sigma} C_k m_k + E\left[\sum_{k=\sigma+1}^{\tau} C_k m_k + V_\tau m_\tau \mathbf{1}_{\tau<\infty} \middle| \mathcal{F}_\sigma\right].
\end{aligned}$$

Since $|\sum_{k=1}^{\sigma} C_k m_k| \leq |\sum_{k=1}^{\infty} C_k m_k| < \infty$ a.s. we can cancel the sum $\sum_{k=1}^{\sigma} C_k m_k$ from both sides. Dividing by m_σ gives the desired result. To prove (i) $\Rightarrow$ (ii) we let $\tau = \infty$ and $\sigma = t$, for $t \in \mathbb{N}$. We are now back to Theorem 2.12 and the proof found there. $\square$

We know from Theorem 2.12 that (ii) in the previous proposition is equivalent to the fact that the value process has the form $V_t = E\left[\sum_{k=t+1}^{\infty} m(t,k) C_k \middle| \mathcal{F}_t\right]$. Thus if we replace the infinite horizon and the time t with two stopping times, we still have the equivalences of Theorem 2.12.

Bibliography

[1] Bühlmann, H. (1976), 'A Probabilistic Approach to Long Term Insurance (Typically Life Insurance)', Lecture given at the *International Congress of Actuaries* in Tokyo

[2] Bühlmann, H. (1992), 'Stochastic discounting', *Insurance: Mathematics and Economics*, 11, 113-127

[3] Campbell, J. Y., Lo, A. W. & MacKinley A. C. (1997), 'The Econometrics of Financial Markets', *Princeton University Press*

[4] Cuthbertson, K. (2000), 'Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange', *John Wiley & Sons*

[5] Duffie, D. (1996), 'Dynamic Asset Pricing Theory', *Princeton University Press*

[6] Duffie, D. (1985) 'Price Operators: Extensions, Potentials, and the Markov Valuation of Securities', *Research Paper No. 813, Graduate School of Business, Stanford University*

[7] Durrett, R. (1996), 'Probability: Theory and Examples, Second edition, *Duxbury Press*

[8] Garman, M. B. (1985), 'Towards a Semigroup Pricing Theory', *The Journal of Finance*, Vol. 40, No. 3, 847-861

[9] Gerber, H. U. (1975), 'A Probabilistic Model for (Life) Contingencies and a Delta–free Approach to Contingency Reserves', *Transactions of the Society of Actuaries*, 28, 127-141

[10] Norberg, R. (2001), 'Financial Mathematics in Life and Pension Insurance', *Lecture notes*

[11] Ottaviani, G. (Ed.) (2000), 'Financial Risk in Insurance', *Springer Verlag*

[12] Papatriandafylou, A. & Waters, R.W. (1984), 'Martingales in Life Insurance', *Scandinavian Actuarial Journal*

[13] Pliska, S.& R. (2000) 'Introduction to Mathematical Finance, Discrete Time Models', *Blackwell Publishers*

[14] Schachermayer, W. (1992), 'Martingale Measures for Discrete Time Processes with Infinite Horizon', Working paper

[15] Williams, D. (1999), 'Probability with Martingales', *Cambridge University Press*

Chapter 3

Cash Flow Valuation in Continuous Time

3.1 Introduction

When an individual or firm is faced with a stream of future cash flows the natural question is what the present value of these cash flows is. The way we value the cash flows is to discount them using some suitable discount rate and then sum them up. If the cash flows and/or the discount rates are stochastic we also take expectations. See Brealey & Myers [4] for the basics on valuation of cash flows, and Copeland et al [7] for an introduction to corporate valuation. In life insurance the prospective reserve is the discounted value of future cash flows. Martin-Löf [21] and Norberg [23] discuss properties of the reserves (prospective and retrospective), and Norberg [22], with applications to insurance in mind, gives an axiomatic approach to valuation. In the approach of no arbitrage pricing it is a well known fact that absence of arbitrage will imply the existence of an equivalent martingale measure under which the expectations are to be taken. The discount rate in this case should then be taken as the risk-free rate. Björk [3] and Duffie [10] are standard text books in this area. If there is no capital market generating the cash flows we can not rely on no arbitrage pricing and we have to choose some probabilities together with a risk-adjusted rate when we value the cash flows. It could even be that different individuals have different perceptions of the probability laws ruling the cash flows and the discount rates. Whatever route we take, the same structure applies: the value is an expected sum of the discounted cash flows. Recently the theory of real options has gained interest in valuation theory. The idea is to identify an embedded option in the investment and add this value to the net present value (calculated as described above). There are two general ways of doing the modelling underlying the real option valuation. Either one models the value directly, or one models the cash flows generating the value and then uses this calculated value as the underlying process in the option valuation. In the latter case we need to understand how the dynamics of the cash flow process influence the dynamics of the value process. We will approach this problem as an application of the Brownian models treated below. In Dixit & Pindyck [9] many examples of the theory of real options are presented, while Copeland & Antikarov [6] focus more on how to apply the theory in practice. Hubalek & Schachermayer [16] discuss the connection of real options to the theory of no arbitrage pricing.

Although one could argue that the cash flows arrive at discrete times, in this paper we choose to work in continuous time. The advantage of this approach is that we can rely on the stochastic calculus of semimartingales, and especially on Itô diffusion models. Consider the following simple model. A firm is facing the (stochastic) cash flow rate $(C_t)_{t\geq 0}$. We define the value at time t as

$$V_t = E_t\left[\int_t^\infty C_s e^{-r(s-t)} ds\right],$$

where $E_t[\cdot]$ denotes that the expectations should be taken with respect to all known information up to time t, and r is some constant discount rate. We rewrite this expression as

$$V_t e^{-rt} = E_t\left[\int_0^\infty C_s e^{-rs} ds\right] - \int_0^t C_s e^{-rs} ds.$$

If we assume that $E\left|\int_0^\infty C_s e^{-rs} ds\right| < \infty$, then this is a decomposition where the discounted present value is the sum of a uniformly integrable martingale and a predictable process. If we denote the martingale by M, we see that

$$d(V_t e^{-rt}) = dM_t - C_t e^{-rt} dt.$$

By using the differential rule for products we can write the dynamics of the present value V as

$$dV_t = (rV_t - C_t)dt + e^{rt} dM_t.$$

We can also make the following heuristic analysis (we follow Cochrane [5]). Define $\Lambda_t = e^{-rt}$ and start again with

$$V_t \Lambda_t = E_t\left[\int_t^\infty C_s \Lambda_s ds\right].$$

We can rewrite this as

$$V_t \Lambda_t = E_t\left[\int_t^{t+h} C_s \Lambda_s ds + V_{t+h}\Lambda_{t+h}\right].$$

Moving $V_t\Lambda_t$ to the right-hand side yields

$$0 = E_t\left[\int_t^{t+h} C_s \Lambda_s ds\right] + E_t\left[V_{t+h}\Lambda_{t+h} - V_t\Lambda_t\right],$$

and by letting $h \downarrow 0$ we can formally write

$$0 = C_t \Lambda_t dt + E_t \left[d(V_t \Lambda_t) \right].$$

The idea when introducing Λ_t is of course to allow for more general discount factors, especially stochastic ones. We also want to generalize the cash flows, allowing them to be modelled as processes of finite variation. The problem of valuation, defined as determining the value process, has connections to forward–backward stochastic differential equations (FBSDE). We show that the general valuation problem, in the Brownian model, can equivalently be written as an FBSDE.

The rest of the paper is organized as follows. In Section 2 we make more precise what we mean by a cash flow process and deflator. Section 3 contains the definition and the basic properties of the valuation process. We show that there exists three equivalent forms on which we can state that the value process is generated by the cash flows and deflator as discussed above. Some extensions are then discussed. Finally, Section 4 contains the case of Brownian models, where we focus on two questions: the connection the valuation problem has to FBSDE's and how the dynamics of the cash flow process and the dynamics of the value process depend on each other. This is then applied to real options.

3.2 Preliminaries

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a complete filtered probability space. The filtration is assumed to satisfy right continuity and $\mathcal{F}_0$ is assumed to contain the P-null sets of $\mathcal{F}$. We let $\mathcal{F}_\infty$ denote the σ-algebra $\bigvee_{t \geq 0} \mathcal{F}_t$, and assume that $\mathcal{F}_\infty = \mathcal{F}$. By an increasing process we mean a process whose paths a.s. are positive, increasing and right continuous. An increasing function has left limits, and thus any increasing process is cadlag (i.e. have paths that are a.s. right continuous and have left limits). We will use the convention $A_{0-} = 0$ a.s. for every increasing process A, but we do not require that $A_0 = 0$ a.s. A process is of finite variation (or an FV process) if it is cadlag and adapted and if almost every sample path is of finite variation on each compact subset of $[0, \infty)$. It is a well known fact that a process is of finite variation if and only if it is the difference between two increasing processes. A process X is said to be optional if the mapping $X : [0, \infty) \times \Omega \to \mathbb{R}$ is measurable when $[0, \infty) \times \Omega$ is given the optional σ-algebra. Since the optional σ-algebra is generated by the family of all adapted process which are cadlag (Elliot [14], Theorem 6.35) every FV process A is optional. If A is an FV process and X is a real valued process on $[0, \infty) \times \Omega$ that is $\mathcal{B} \times \mathcal{F}$-measurable (here $\mathcal{B}$ denotes the Borel σ-algebra on $[0, \infty)$) we define the Stieltjes integral

$$(X \cdot A)_t(\omega) = \int_{[0,t]} X_s(\omega) dA_s(\omega), \tag{3.1}$$

whenever it exists. If $X \cdot A$ exists for all $t \in [0, \infty)$ and almost all $\omega \in \Omega$ then $(X \cdot A)_t$ defines a process with $(X \cdot A)_0 = X_0 A_0$. If X is optional there exists an optional version of $(X \cdot A)$. If A is an FV process then $|A|_t(\omega) = \int_0^t |dA_s(\omega)|$ denotes the total variation of A. $|A|$ is adapted and cadlag and $|A|_0 = |A_0|$. If A is an FV process and X is a measurable process then the integral $\int_0^t X_s(\omega)|dA_s(\omega)|$ denotes integration with respect to $d|A|$. A semimartingale is an adapted and cadlag process $(X_t)_{t\geq 0}$ having a decomposition $X_t = X_0 + M_t + A_t$, where M is a local martingale and A is an FV process fulfilling $M_0 = 0$ a.s. and $A_0 = 0$ a.s.

3.2.1 Cash flows and deflators

Definition 3.1 *A* cash flow process $(C_t)_{t\geq 0}$ *is an FV process.*

This definition of the cash flow process makes it trivially a semimartingale. We can also use it as an integrator, thus making it possible to define processes of the type as in Equation (3.1).

Definition 3.2 *A* deflator *is a strictly positive semimartingale that is finite a.s.*

This is a generalization of the definition in Duffie [10], who defines a deflator to be a strictly positive Itô process. The reason for demanding the deflator to be a semimartingale, and not a more general process, is that we want to use the differentiation rule valid for semimartingales. We note here that if Λ is a deflator then both $1/\Lambda$ and $\ln \Lambda$ are well defined, and since $1/x$ and $\ln x$ are twice continuously differentiable on $(0, \infty)$, both $1/\Lambda$ and $\ln \Lambda$ are semimartingales (this follows from Theorem 32 of Chapter II in Protter [25]).

Definition 3.3 *Given a deflator Λ, the* discount process *implied by Λ is defined by*

$$m(s,t) = \frac{\Lambda(t)}{\Lambda(s)}, \quad s,t \geq 0.$$

The following proposition, whose proof is an immediate consequence of the definition of deflator, presents some important properties of the discount process. In Armerin [1] the properties of m proved in the following proposition were taken as the definition of the discount process. The reason for this exchange of definition and proposition is that the definition given in Armerin [1] is the more natural one and works well in discrete time. In continuous time, however, it is easier to work with the deflator as the defining object.

Proposition 3.4 *Let m be a discount process implied by the deflator Λ. Then*

(i) $m(s,t)$ *is* $\mathcal{F}_{\max(s,t)}$*-measurable for every* $s,t \in [0,\infty)$.

(ii) $0 < m(s,t) < \infty$ *a.s. for every* $s,t \in [0,\infty)$.

(iii) $m(s,t) = m(s,u)m(u,t)$ *a.s for every* $0 \leq s \leq u \leq t$.

A discount process fulfilling $0 < m(s,t) \leq 1$ a.s. for every $0 \leq s \leq t < \infty$ will be referred to as a *normal* discount process. We see that m is normal if and only if Λ is nondecreasing. We close this section with the observation that a discount process m, with deflator Λ, can be written in the form

$$m(s,t) = \exp\left(-\int_s^t \lambda(u)du\right) \tag{3.2}$$

if and only if $\ln \Lambda(t)$ is a.s. absolutely continuous in t with density $-\lambda(t,\omega)$.

3.3 Valuation

Definition 3.5 *Given a cash flow process C and a deflator Λ such that*

$$E\left[\int_{[0,\infty)} \Lambda_s |dC_s|\right] < \infty,$$

the value process *is defined for $t \in [0,\infty)$ as*

$$V_t = \frac{1}{\Lambda_t} E\left[\int_{(t,\infty)} \Lambda_s dC_s \middle| \mathcal{F}_t\right].$$

By noting that $\int_{[0,\infty)} = \int_{[0,t]} + \int_{(t,\infty)}$ and using the fact that every optional process is adapted (Jacod & Shiryaev [17], Proposition 1.21) we get

$$V_t\Lambda_t = E\left[\int_{[0,\infty)} \Lambda_s dC_s \middle| \mathcal{F}_t\right] - \int_{[0,t]} \Lambda_s dC_s = M_t - (\Lambda \cdot C)_t. \tag{3.3}$$

Since

$$E|M_t| = E\left|E\left[\int_{[0,\infty)} \Lambda_s dC_s \middle| \mathcal{F}_t\right]\right| \leq E\left[\int_{[0,\infty)} \Lambda_s |dC_s|\right] < \infty$$

for every $t \in [0,\infty)$, $M_t = E\left[\int_{[0,\infty)} \Lambda_s dC_s \middle| \mathcal{F}_t\right]$ is a uniformly integrable martingale. The filtration $(\mathcal{F}_t)$ is right continuous, thus there exists a modification of M that is right continuous. We also remark here that $M_\infty = \lim_{t\to\infty} M_t = \int_0^\infty \Lambda dC_s$ a.s. These facts immediately follow from Elliot [14], Theorem 4.11. Now, since both M and $\Lambda \cdot C$ are right continuous and adapted the value process is also right continuous and adapted. From this it follows that the value process is optional. Equation (3.3) implies that

$$V_t = \frac{M_t}{\Lambda_t} - \frac{(\Lambda \cdot C)_t}{\Lambda_t}.$$

Since M is a (true) martingale it is especially a semimartingale. $C \cdot \Lambda$ is a process of finite variation, and is thus also a semimartingale. Since Λ is a strictly positive semimartingale it follows that $1/\Lambda$ is a strictly positive semimartingale, and since the product of two semimartingales is again a semimartingale we see finally that V is a semimartingale. We remark here the fact that Delbaen & Schachermayer [8] show that it is reasonable to model the price process of an financial asset as a semimartingale. Norberg [23] defines the prospective reserve of a life insurance company in the same way as we have defined the value here. In life insurance the value at time t of a cash flow stream is defined as, using our definitions, $E\left[\frac{1}{\Lambda_t}\int_{[0,\infty)}\Lambda_s dC_s \middle| \mathcal{F}_t\right]$. For more on the reserves in life insurance see Norberg [23] and references therein and Norberg [24]. As in the discrete time case (see Armerin [1]) there exist three equivalent representations of the value process. The only general assumptions made here are measurability conditions on C and Λ and the integrability condition making M into a uniformly integrable martingale. We also need a condition essentially stating that the discounted value goes to zero as t tends to infinity (see the comment in Armerin [1] on rational bubbles what happens if we discard this condition).

Theorem 3.6 *Let C and Λ be a cash flow process and a deflator respectively, such that $E\left[\int_{[0,\infty)}\Lambda_s |dC_s|\right] < \infty$. Then the following three statements are equivalent.*

(i) For every $t \in [0,\infty)$

$$V_t = \frac{1}{\Lambda_t} E\left[\int_{(t,\infty)} \Lambda_s dC_s \middle| \mathcal{F}_t\right]. \tag{3.4}$$

(ii) (a) For every $t \in [0,\infty)$

$$M_t = V_t \Lambda_t + \int_{[0,t]} \Lambda_s dC_s \tag{3.5}$$

is a uniformly integrable martingale, and

(b) $V_t\Lambda_t \to 0$ a.s. when $t \to \infty$.

(iii) For each $t \in [0,\infty)$ we have

(a) For every $h > 0$

$$V_t\Lambda_t = E\left[V_{t+h}\Lambda_{t+h} + \int_{(t,t+h]} \Lambda_s dC_s \middle| \mathcal{F}_t\right], \tag{3.6}$$

and

(b) $\lim_{T\to\infty} E\left[V_{t+T}\Lambda_{t+T} | \mathcal{F}_t\right] = 0$.

Proof. We will show (i)⇔(ii) and (i)⇔(iii).

(i)⇔(ii): To prove the 'only if' part we rewrite Equation (3.4) as $M_t = V_t\Lambda_t +$

$\int_{[0,t]} \Lambda_s dC_s$, $t \in [0, \infty)$. We know from above that M is a uniformly integrable martingale, and using this together with the fact that $M_t \xrightarrow{a.s.} M_\infty = \int_0^\infty \Lambda_s dC_s$ gives

$$\lim_{t\to\infty} V_t \Lambda_t = \lim_{t\to\infty} \left(M_t - \int_{[0,t]} \Lambda_s dC_s \right) = 0 \text{ a.s.}$$

Turning to the 'if' part we let $t \to \infty$ in Equation (3.5). Using (ii) (b) we get $M_\infty = \int_{[0,\infty)} \Lambda dC_s$. Taking the conditional expectation with respect to the σ-algebra $\mathfrak{F}_t$ we get

$$\begin{aligned} V_t\Lambda_t + \int_{[0,t]} \Lambda_s dC_s &= M_t = E\left[M_\infty | \mathfrak{F}_t\right] \\ &= E\left[\int_{[0,\infty)} \Lambda_s dC_s \middle| \mathfrak{F}_t \right] \\ &= \int_{[0,t]} \Lambda_s dC_s + E\left[\int_{(t,\infty)} \Lambda_s dC_s \middle| \mathfrak{F}_t \right]. \end{aligned}$$

Rearranging this relation gives the desired result.

(i)⇔(iii): For the 'if' part take $h > 0$. We get, using Equation (3.4),

$$\begin{aligned} V_t\Lambda_t &= E\left[\int_{(t,\infty)} \Lambda_s dC_s \middle| \mathfrak{F}_t \right] \\ &= E\left[\int_{(t,t+h]} \Lambda_s dC_s + \int_{(t+h,\infty)} \Lambda_s dC_s \middle| \mathfrak{F}_t \right] \\ &= E\left[\int_{(t,t+h]} \Lambda_s dC_s + E\left[\int_{(t+h,\infty)} \Lambda_s dC_s \middle| \mathfrak{F}_{t+h} \right] \middle| \mathfrak{F}_t \right] \\ &= E\left[\int_{(t,t+h]} \Lambda_s dC_s + \Lambda_{t+h} V_{t+h} \middle| \mathfrak{F}_t \right]. \end{aligned}$$

Since

$$\left| \int_{(t,\infty)} \Lambda_s dC_s \right| \le \int_{(t,\infty)} \Lambda_s |dC_s| \le \int_{[0,\infty)} \Lambda_s |dC_s|$$

and $\int_{[0,\infty)} \Lambda_s |dC_s|$ is integrable, we use the Dominated Convergence Theorem to get for every $A \in \mathfrak{F}_t$

$$\lim_{T\to\infty} E\left[V_{t+T}\Lambda_{t+T}\mathbf{1}_A\right] = \lim_{T\to\infty} E\left[\int_{(t+T,\infty)} \Lambda_s dC_s \mathbf{1}_A \right] = E\left[\lim_{T\to\infty} \int_{(t+T,\infty)} \Lambda_s dC_s \mathbf{1}_A \right] = 0,$$

where the last equality follows from the fact that $\int_{[0,\infty)} \Lambda_s dC_s$ is finite a.s. To prove the other direction of the equivalence we let $T \to \infty$ in Equation (3.6):

$$\begin{aligned} V_t \Lambda_t &= \lim_{T\to\infty} E\left[V_{t+T}\Lambda_{t+T}|\mathcal{F}_t\right] + \lim_{T\to\infty} E\left[\int_{(t,t+T]} \Lambda_s dC_s \middle| \mathcal{F}_t\right] \\ &= \lim_{T\to\infty} E\left[\int_{(t,t+T]} \Lambda_s dC_s \middle| \mathcal{F}_t\right]. \end{aligned}$$

Again we used the Dominated Convergence Theorem to interchange the limit and the expectation to get the desired conclusion. □

3.3.1 Stopping Times

It is not difficult to see that Theorem 3.6 can be generalized to allow also for stopping times. The content of the following theorem is that we can strengthen the results of Theorem 3.6 by replacing both the infinite horizon and the time of valuation with a stopping time. For the proof we essentially only need to use the Theorem of Optional Stopping for uniformly integrable martingales.

Theorem 3.7 *Let C and Λ be a cash flow process and a deflator respectively such that $E\left[\int_{[0,\infty)} \Lambda_s|dC_s|\right] < \infty$. Then the following two statements are equivalent.*

(i) For all stopping times σ and τ such that $0 \le \sigma \le \tau$ a.s.

$$V_\sigma = \frac{1}{\Lambda_\sigma} E\left[V_\tau \mathbf{1}_{\tau<\infty} + \int_{(\sigma,\tau]} \Lambda_s dC_s \middle| \mathcal{F}_\sigma\right] \quad \text{on } \{\sigma < \infty\}.$$

(ii) (a) For every $t \in [0,\infty)$

$$M_t = V_t\Lambda_t + \int_{[0,t]} \Lambda_s dC_s$$

is a uniformly integrable martingale, and

(b) $V_t\Lambda_t \to 0$ a.s. when $t \to \infty$.

Proof. We first show (ii)⇒(i). Take $n \in \mathbb{N}$. Then

$$\begin{aligned} M_{\tau\wedge n} &= V_{\tau\wedge n}\Lambda_{\tau\wedge n} + \int_{[0,\tau\wedge n]} \Lambda_s dC_s \\ &\xrightarrow{a.s.} V_\tau\Lambda_\tau\mathbf{1}_{\tau<\infty} + \int_{[0,\tau]} \Lambda_s dC_s \text{ as } n \to \infty. \end{aligned}$$

From this and the Theorem of Optional Stopping we get

$$V_\sigma \Lambda_\sigma \mathbf{1}_{\sigma<\infty} + \int_{[0,\sigma]} \Lambda_s dC_s = M_\sigma = E\left[M_\tau | \mathcal{F}_\sigma\right] = E\left[V_\tau \Lambda_\tau \mathbf{1}_{\tau<\infty} + \int_{[0,\tau]} \Lambda_s dC_s \middle| \mathcal{F}_\sigma \right].$$

Using the fact that $\int_{[0,\sigma]} \Lambda_s dC_s$ is finite a.s. and measurable with respect to $\mathcal{F}_\sigma$ yields the desired result. To show (i)⇒(ii) we let $\tau = \infty$ and $\sigma = t$ and then use the proof of Theorem 3.6. □

3.3.2 On the local dynamics of the value process

In this section we will comment on the local behavior of the value process. The starting point is relation (3.3):

$$V_t \Lambda_t = M_t - (\Lambda \cdot C)_t.$$

Since all the processes in this expression are semimartingales we can use the differentiation rule for products of semimartingales (Protter [25], Chapter II.6, Corollary 2) to get

$$d(V_t \Lambda_t) = V_{t-} d\Lambda_t + \Lambda_{t-} dV_t + d[V, \Lambda]_t = dM_t - \Lambda_t dC_t. \tag{3.7}$$

By dividing this expression with Λ_{t-} and using the facts that $\Lambda_t dC_t = \Lambda_{t-} dC_t + \Delta\Lambda_t \Delta C_t$ we get

$$dV_t = V_{t-}\left(-\frac{d\Lambda_t}{\Lambda_{t-}}\right) - dC_t - \frac{1}{\Lambda_{t-}}\big(d[V, \Lambda]_t + \Delta\Lambda_t \Delta C_t\big) + \frac{1}{\Lambda_{t-}} dM_t.$$

By introducing

$$dR_t = -\frac{d\Lambda_t}{\Lambda_{t-}} \text{ and } N_t = \int_0^t \frac{1}{\Lambda_{s-}} dM_s$$

we can write

$$dV_t = V_{t-} dR_t - dC_t - \frac{1}{\Lambda_{t-}}\big(d[V, \Lambda]_t + \Delta\Lambda_t \Delta C_t\big) + dN_t.$$

If $\Lambda_0 = 1$, then Λ is the stochastic exponential of R. Hence, we can interpret R as a discount rate associated with Λ. To improve the economical interpretation of Equation (3.7) note that if we have a cash flow given by $r_t^f dt$ for $t \in [0, \infty)$, where r_t^f is measurable and adapted and such that for a.e. $\omega \in \Omega$ we have $0 \leq r_t^f(\omega)$ for every $t \in [0, \infty)$, and if the value process of this cash flow stream fulfills $V_t \equiv 1$, then we can think of r_t^f as a locally risk-free interest rate (see Cochran [5] p. 29 ff.). Inserting these specifications into Equation (3.7) yields the relation

$$r_t^f dt = -\frac{d\Lambda_t}{\Lambda_t} + \frac{1}{\Lambda_t} d\widetilde{M}_t,$$

where $\widetilde{M} = E\left[\int_0^\infty r_s^f \Lambda_s ds | \mathcal{F}_t\right]$. We can write Equation (3.3) as

$$\Lambda_t = \widetilde{M}_t - \int_0^t r_s^f \Lambda_s ds. \tag{3.8}$$

The fact that $r_t^f \geq 0$ a.s. implies that $A_t = \int_0^t r_s^f \Lambda_s ds$ is an increasing process. This in turn means that Λ is a non-negative supermartingale. An adapted and cadlag process is a *potential* if it is a non-negative supermartingale that tends to 0 a.s. as time goes to infinity. Since

$$E\left[\Lambda_t\right] = E\left[\int_0^\infty r_s^f \Lambda_s ds\right] - \int_0^t E\left[r_s^f \Lambda_s\right] ds \to 0 \text{ a.s. as } t \to \infty$$

we have proved the following proposition.

Proposition 3.8 *If there exists a risk-less asset, then the deflator pricing this asset is a potential.*

An example where the fact that the deflator is a potential is crucial is given in Rogers [26]. Let us return to Equation (3.8). This can be seen as an equation for Λ. If the filtration $(\mathcal{F}_t)$ is generated by a Brownian motion, then we know that every square integrable martingale can be written as an Itô integral. In this case we only need one more asset priced by Λ to determine Λ (see Harrison & Kreps [15]). Now, if the deflator Λ assigning the cash flow stream given by $r_t^f dt$ the value 1 for all $t \in [0, \infty)$, then, when using the same Λ for valuing another cash flow stream C, we can express the differential of V in terms of the rate r_t^f. If we assume that Λ is a continuous process, then Equation (3.7) can be written

$$V_{t-} d\Lambda_t + \Lambda_t dV_t + d[V, \Lambda]_t^c = dM_t - \Lambda_t dC_t,$$

or

$$dV_t = \left(-\frac{d\Lambda_t}{\Lambda_t}\right) V_{t-} - dC_t - \frac{1}{\Lambda_t} d[V, \Lambda]_t^c + \frac{1}{\Lambda_t} dM_t.$$

Replacing $d\Lambda_t$ by $-r_t^f \Lambda_t dt + d\widetilde{M}_t$ we can write this as

$$dV_t + dC_t = r_t^f V_t dt - \frac{1}{\Lambda_t} d[V, \Lambda]_t^c + \frac{1}{\Lambda_t} dM_t - \frac{V_t}{\Lambda_t} d\widetilde{M}_t.$$

Taking conditional expectations we can write this formally as

$$E\left[dV_t + dC_t | \mathcal{F}_t\right] = r_t^f V_t dt - \frac{1}{\Lambda_t} d[V, \Lambda]_t^c.$$

If we further make the assumption that V is a strictly positive process and write $d[V, \Lambda]_t^c = dV_t d\Lambda_t$, we arrive at

$$E\left[\left.\frac{dV_t + dC_t}{V_t}\right| \mathcal{F}_t\right] = r_t^f dt - E\left[\left.\frac{dV_t}{V_t} \frac{d\Lambda_t}{\Lambda_t}\right| \mathcal{F}_t\right].$$

The left hand side of this equation is the instantaneous net return of the value process at t. We have now decomposed the expected return of the value process into two parts: the risk-free part ($r_t^f dt$) and a risk premium $\left(-E\left[\frac{dV_t}{V_t}\frac{d\Lambda_t}{\Lambda_t}\,\middle|\,\mathcal{F}_t\right]\right)$. Thus, if $\frac{dV_t}{V_t}$ and $\frac{d\Lambda_t}{\Lambda_t}$ are negatively correlated there is a positive risk premium, and if they are positively correlated the risk premium is negative. The intuition is that a risky investment is desirable if its value is high in 'bad' states of the world (when we really need money) and low in 'good' states of the economy (when everything else is good). An investment with such properties will have a high price (since demand for this desirable investment opportunity is high), and thus a low expected return. This allows for the interpretation of $d\Lambda_t/\Lambda_t$ as a measure of how 'bad' a state of the economy is. See Cochran ([5] Section 1.5 and Part III) for more on this type of asset pricing in continuous time.

We will now solve Equation (3.8) when the risk-less rate is equal to the constant $r_f > 0$. Inserting this in Equation (3.4) gives the relation

$$\Lambda_t = E\left[\int_{(t,\infty)} r_f \Lambda_s ds \,\middle|\, \mathcal{F}_t\right]. \tag{3.9}$$

This is a stochastic differential equation for the discount factor Λ. Inserting the function $\Lambda_s = e^{-r_f s}$, on the right-hand side of Equation (3.9) gives

$$E\left[\int_t^\infty r_f e^{-r_f s} ds \,\middle|\, \mathcal{F}_t\right] = \left[-e^{-r_f s}\right]_t^\infty = e^{-r_f t} = \Lambda_t.$$

Hence this Λ is a solution to Equation (3.9). If $K \in \mathbb{R}$, then also $Ke^{-r_f s}$ is a solution. Are there any others? The answer to this question is yes, as is seen on the following proposition.

Proposition 3.9 *Let $r_f > 0$ be a given real number and consider the equation*

$$X_t = E\left[\int_{(t,\infty)} r_f X_s ds \,\middle|\, \mathcal{F}_t\right]. \tag{$\star$}$$

We say that Λ is a solution to ($\star$) if Λ fulfills this relation and the integrability condition

$$E\left[\left|\int_0^\infty r_f \Lambda_s ds\right|\right] < \infty.$$

Λ is a solution to ($\star$) if and only if it can be written as

$$\Lambda_t = e^{-r_f t}\left(\Lambda_0 + \int_0^t e^{r_f s} dM_s\right) \tag{3.10}$$

for some uniformly integrable martingale M.

We need the following lemma due to Lépingle [19].

Lemma 3.10 *Let N be a local martingale that is cadlag and fulfills $N_0 = 0$ a.s., and let U be a predictable increasing process with $U_0 > 0$ a.s. If $\int_0^t (1/U_s) dN_s$ converges a.s. as $t \to \infty$ and the limit is finite, then*

$$\frac{N_t}{U_t} \to 0 \text{ a.s. as } t \to \infty \text{ on } \{U_\infty = \infty\}.$$

We now turn the proof of Proposition 3.9.

Proof. Assume a Λ on the form given in Equation (3.10) and define the local martingale N according to

$$N_t = \int_0^t e^{r_f s} dM_s.$$

With this N and $U_t = e^{r_f t}$ we have

$$\int_0^t \frac{1}{U_s} dN_s = M_t \xrightarrow{a.s.} M_\infty \in L^1.$$

Since $P(U_\infty = \infty) = 1$ we can use Lemma 3.10 to conclude that

$$e^{-r_f t} N_t \to 0 \ a.s.$$

Using the integration by parts formula

$$d(e^{-r_f t}(\Lambda_0 + N_t)) = -r_f e^{-r_f t}(\Lambda_0 + N_t)dt + e^{-r_f t} e^{r_f t} dM_t = -r_f e^{-r_f t}(\Lambda_0 + N_t)dt + dM_t$$

we see that for $0 \le t \le T$ we have

$$\int_t^T r_f e^{-r_f s}(\Lambda_0 + N_s)ds = M_T - M_t - e^{-r_f T}(\Lambda_0 + N_T) + e^{-r_f t}(\Lambda_0 + N_t).$$

Since $e^{-r_f T}(\Lambda_0 + N_T) \to 0$ a.s. as $T \to \infty$, we arrive at (since M by assumption is a uniformly integrable martingale it will converge a.s. to an integrable random variable M_∞)

$$e^{-r_f t}(\Lambda_0 + N_t) = \int_t^\infty r_f e^{-r_f s}(\Lambda_0 + N_s)ds - M_\infty + M_t.$$

Taking expectations with respect to $\mathcal{F}_t$ gives

$$e^{-r_f t}(\Lambda_0 + N_t) = E\left[\int_t^\infty r_f e^{-r_f s}(\Lambda_0 + N_s)ds \,\middle|\, \mathcal{F}_t\right],$$

or

$$\Lambda_t = E\left[\int_t^\infty r_f \Lambda_s ds \,\middle|\, \mathcal{F}_t\right].$$

By letting $t = 0$ we get

$$E\left[\left|\int_0^\infty r_f e^{r_f s}(\Lambda_0 + N_s)ds\right|\right] \leq |\Lambda_0| + E\left[|M_\infty|\right] + |M_0| < \infty,$$

and we can conclude that $e^{-r_f t}(\Lambda_0 + N_t)$ is a solution. Now assume that Λ is a solution. We rewrite Equation (3.9) as

$$\Lambda_t = -\int_0^t r_f \Lambda_s ds + E\left[\int_0^\infty r_f \Lambda_s ds \middle| \mathcal{F}_t\right]$$

and introduce the UI martingale

$$M_t^\Lambda = E\left[\int_0^\infty r_f \Lambda_s ds \middle| \mathcal{F}_t\right].$$

We get

$$\begin{aligned}\Lambda_t e^{r_f t} &= \Lambda_0 + \int_0^t r_f \Lambda_s e^{r_f s} ds + \int_0^t e^{r_f s} dM_s^\Lambda - \int_0^t r_f \Lambda_s e^{r_f s} ds \\ &= \Lambda_0 + \int_0^t e^{r_f s} dM_s^\Lambda.\end{aligned}$$

Hence, Λ has the desired form. □

3.4 Brownian models

We will from now on assume that the cash flow process and deflator both are driven by a (possibly multi-dimensional) Brownian motion. The model we use consists of a time-homogeneous Itô diffusion representing some state(s) that influence the cash flows and the discount factor. Let B be an n-dimensional Brownian motion on our probability space. We will let $(\mathcal{F}_t)$ denote the standard Brownian filtration generated by B augmented with all P-null sets of $\mathcal{F}$.

3.4.1 The value process as a solution to an FBSDE

The aim of this section is to show the close connection between the value process and a class of forward–backward stochastic differential equations (FBSDE). We begin by looking at backward stochastic differential equations (BSDE). Consider the problem of finding adapted solutions to the equation

$$\begin{cases} dY_t &= 0, \ 0 \leq t < T, \\ Y_T &= \xi, \end{cases}$$

where $T > 0$ is a fixed time and $\xi \in L^2(\Omega, \mathcal{F}_T)$. $Y_t = \xi$, $0 \leq t \leq T$, satisfies the equation but it is not adapted to the filtration. The key observation to continue is that $Y_t = E[\xi|\mathcal{F}_t]$, $0 \le t \le T$, satisfies the terminal condition and is adapted to$(\mathcal{F}_t)$. Since ξ is square integrable, $Y_t = E[\xi|\mathcal{F}_t]$ is a martingale and can be represented as $Y_t = Y_0 + \int_0^t Z_s dB_s$ for some a.s. unique adapted and square integrable process Z. This Y satisfies for $0 \le t \le T$

$$\begin{cases} dY_t &= Z_t dB_t \\ Y_T &= \xi. \end{cases}$$

We now define a solution to this problem as a pair (Y, Z) of adapted processes. We have thus been able to find an adapted solution, not to our original problem, but to a similar one. It has been shown that this is the 'right' way to do it, see Ma & Yong [20] Chapter 1. The general BSDE is

$$\begin{cases} dY_t &= -f(t, Y_t, Z_t)dt + Z_t dB_t, \\ Y_T &= \xi, \end{cases}$$

where B is a d-dimensional Brownian motion, $\xi \in L^2(\Omega, \mathcal{F}_T)$ and $T > 0$ is a fixed time. It has turned out that there is a variety of problems that be formulated in the context of BSDE; see e.g. Ma & Yong [20] and references therein for the theory and applications, and El Karoui et al [13] and the contribution of El Karoui & Quenez in [2] for applications to finance. The extension to forward-backward stochastic differential equations (FBSDE) is done by introducing another state variable X moving 'forward':

$$\begin{cases} dX_t &= b(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dB_t, \\ dY_t &= h(t, X_t, Y_t, Z_t)dt + Z_t dB_t, \\ X(0) &= x, \\ Y(T) &= g(X(T)). \end{cases}$$

For the technical assumptions on b, σ and h see Ma & Yong [20]. We will be interested in the case when the time horizon is infinite. There is no immediate generalization of the above equation for this case, but Ma & Yong propose the additional requirement that Y be bounded a.s. uniformly in $t \in [0, \infty)$. Duffie et al [11] uses an FBSDE with infinite horizon to solve Black's consol rate conjecture. We will now show that there is an equivalent formulation of the definition of the value process in the form of an FBSDE. Let $b : [0, \infty) \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ and $\sigma : [0, \infty) \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ be Borel measurable functions. Consider the following problems:

(P1) Find a pair of adapted, locally square integrable processes (X, V) such that for $t \in [0, \infty)$

$$\begin{cases} dX_t &= b(t,X_t,V_t)dt + \sigma(t,X_t,V_t)dB_t \\ V_t &= E\left[\int_t^\infty \exp\left(-\int_t^s \lambda(u,X_u,V_u)du\right) g(s,X_s,V_s)ds \middle| \mathcal{F}_t\right] \\ X_0 &= x, \end{cases}$$

(P2) Find a triplet (X, V, Z) of adapted process such that

$$\begin{cases} dX_t &= b(t,X_t,V_t)dt + \sigma(t,X_t,V_t)dB_t \\ dV_t &= \left[\lambda(t,X_t,V_t)V_t - g(t,X_t,V_t)\right]dt + Z_t dB_t \\ X_0 &= x, \\ E\left[V_T \exp\left(-\int_0^T \lambda(s,X_s,V_s)ds\right) \middle| \mathcal{F}_t\right] &\to 0 \text{ a.s. as } T \to \infty \text{ for every } t \in [0,\infty). \end{cases}$$

Here X is the external process influencing the cash flows and the deflator:

$$\begin{aligned} dC_t &= g(t,X_t,V_t)dt \\ d\Lambda_t &= -\lambda(t,X_t,V_t)\Lambda_t dt, \ \Lambda_0 = 1. \end{aligned}$$

The process X could be macro economical (e.g. inflation, GDP or some exchange rate) or it could be a firm specific variable (e.g. the level of knowledge among the workers of the firm or a measure of progress in the R&D department). We now precise what we mean by an adapted solution to the FBSDE **(P2)**. To begin with we let $L^2(C([0,T];\mathbb{R}^n))$ denote the set of $(\mathcal{F}_t)$-progressively measurable continuous processes X taking values in $\mathbb{R}^n$ such that $E\left[\sup_{t\in[0,T]} \|X(t)\|^2\right] < \infty$, and let $L^2(0,T;\mathbb{R}^n)$ denote the set of $(\mathcal{F}_t)$-progressively measurable processes X taking values in $\mathbb{R}^n$ and such that $\int_0^T E\left[\|X(t)\|^2\right] dt < \infty$. Following Ma & Yong [20] we say that (X, Y, Z) is an adapted solution to **(P2)** if $(X,Y,Z)\big|_{[0,T]} \in L^2(C([0,T];\mathbb{R}^d)) \times L^2(C([0,T];\mathbb{R})) \times L^2(0,T;\mathbb{R}^n)$. The following theorem shows the equivalence between **(P1)** and **(P2)**. It is a generalization of Theorem 3.1 in Chapter 8 in Ma & Yong [20].

Theorem 3.11 *Assume that*

(i) For every $x, y \in \mathbb{R}$

$$0 \le \lambda(t,x,y) \le h(t),$$

where $h : \mathbb{R}_+ \to \mathbb{R}$ *is such that* $\exp\left(\int_0^t h(s)ds\right) < \infty$ *for every* $t \ge 0$.

(ii) The function g *fulfills for every* $x, y \in \mathbb{R}$

$$0 \le g(t,x,y) \le f(t),$$

where $\int_0^\infty f(t)dt < \infty$.

*Under these assumptions if (X, V, Z) is an adapted solution to **(P2)**, then (X, V) is an adapted solution to **(P1)**. Conversely, if (X, V) is an adapted solution to **(P1)**, then there exists an adapted, $\mathbb{R}^n$-valued square integrable process Z such that (X, V, Z) is an adapted solution to **(P2)**.*

Proof. To prove the first statement we assume that (X, V, Z) is an adapted solution to **(P2)** and fix a $t > 0$. Using the integration by parts formula and the property of the solution to **(P2)** we get for every $T \geq t$

$$V_T \Lambda_T = V_t \Lambda_t - \int_t^T \Lambda_s g_s ds + \int_t^T \Lambda_s Z_s dB_s. \tag{3.11}$$

We note that the process $(\Lambda_t Z_t)_{t \in [0,\infty)}$ is measurable and adapted, and fulfills

$$\int_0^T E\left[\Lambda_t^2 Z_t^2\right] dt \leq \int_0^T E\left[Z_t^2\right] dt < \infty$$

(since condition (i) implies $\Lambda_t \leq 1$). Taking conditional expectations with respect to $\mathcal{F}_t$ of Eq. (3.11) we get

$$E\left[V_T \Lambda_T | \mathcal{F}_t\right] = V_t \Lambda_t - E\left[\int_t^T \Lambda_s g_s ds \middle| \mathcal{F}_t\right]. \tag{3.12}$$

Letting $T \to \infty$ in Eq. (3.12) and using the Monotone Convergence Theorem yields the desired conclusion.

For the other direction assume that (X, V) is an adapted solution to problem **(P1)**. We get

$$V_t \Lambda_t = E\left[\int_0^\infty \Lambda_s g_s ds \middle| \mathcal{F}_t\right] - \int_0^t \Lambda_s g_s ds. \tag{3.13}$$

From assumptions (i) and (ii) it follows that

$$\left(\int_0^\infty \Lambda_t g_t dt\right)^2 \leq \left(\int_0^\infty g_t dt\right)^2 \leq \left(\int_0^\infty f(t) dt\right)^2 < \infty.$$

Thus the first processes on the right of Eq. (3.13) is a square integrable martingale. Defining it as M we thus have $V_t = M_t/\Lambda_t - \int_0^t \Lambda_s g_s ds/\Lambda_t$. Integrating by parts yields

$$\begin{aligned} dV_t &= \frac{1}{\Lambda_t} dM_t - \frac{M}{\Lambda_t^2} d\Lambda_t - \frac{\Lambda_t g_t}{\Lambda_t} dt + \frac{1}{\Lambda_t^2}\left(\int_0^t \Lambda_s g_s ds\right) d\Lambda_t \\ &= \left\{\lambda_t V_t - g_t\right\} dt + \frac{1}{\Lambda_t} dM_t. \end{aligned}$$

The Martingale Representation Theorem (see e.g. Theorem 4.3.4 in Øksendal [27]) implies that there exists an a.s. unique stochastic process $\varphi(s, \omega)$ such that for every $t \in [0,\infty)$ $M_t = M_0 + \int_0^t \varphi(s,\omega)dB_s$ and $E\left[\int_0^t \varphi^2(s,\omega)ds\right] < \infty$. Defining $Z_t = \varphi_t/\Lambda_t$ we see that

$$dV_t = \{\lambda_t V_t - g_t\}\, dt + Z_t dB_t.$$

Since X and V are locally square integrable and adapted, and using condition (ii) we get

$$\int_0^T E\left[Z_t^2\right] dt \le \exp\left(\int_0^T h(s)ds\right) \int_0^T E\left[\varphi_t^2\right] dt < \infty$$

for every $T > 0$, (X, V, Z) is an adapted solution to **(P2)**. We only have to check that $E\left[V_T\Lambda_T|\mathcal{F}_t\right] \to 0$ a.s. For this purpose fix a $T > t$. The relation $V_t\Lambda_t = M_t - \int_0^t \Lambda_s g_s ds$ gives

$$E\left[V_T\Lambda_T|\mathcal{F}_t\right] = V_t\Lambda_t - E\left[\int_t^T \Lambda_s g_s ds \,\middle|\, \mathcal{F}_t\right].$$

Letting $T \to \infty$ and again using the Theorem of Monotone Convergence, the right hand side converges to 0 a.s. □

3.4.2 Recovering cash flows from their value process

It is obvious that a given cash flow process uniquely determines a value process. In this section we want to answer the opposite question: given a value process V, does there always exist a cash flow process generating this process, and if it exists, is it unique? To simplify we will make the following assumption.

Assumption 3.12 *The instantaneous rate is constant, i.e. the deflator is given by $\Lambda_t = e^{-\lambda t}$ for some $\lambda \in \mathbb{R}$, and the cash flow process is assumed to be absolutely continuous: $dC_t = c_t dt$. The process c is further assumed to be an Itô diffusion such that $E\left[\left(\int_0^\infty e^{-\lambda t}c_t dt\right)^2\right] < \infty$.*

Given these assumptions, $M_t = E\left[\int_0^\infty e^{-\lambda s}c_s ds \,\middle|\, \mathcal{F}_t\right]$ is a square integrable martingale and from the Martingale Representation Theorem we know that there exists a measurable and adapted process Z fulfilling $E\left[\int_0^t Z_s^2 ds\right] < \infty$ for $t \in [0,\infty)$ such that $M_t = M_0 + \int_0^t Z_s dB_s$. The dynamics of the value process V given by Equation (3.7) is under these assumptions

$$dV_t = (\lambda V_t - c_t)dt + e^{\lambda t} Z_t dB_t.$$

Let $\mu(x)$ and $\sigma(x)$ be two functions such that $dV_t = \mu(V_t)dt + \sigma(V_t)dB_t$ possesses a strong solution. Now, given μ and σ we see that in order for the value process dynamics to be consistent with the dynamics of the value process, in terms of the cash flow process we must have

$$\begin{cases} \mu(V_t) &= \lambda V_t - c_t \\ \sigma(V_t) &= Z_t e^{\lambda t}. \end{cases}$$

Thus the cash flow process generating this value process must fulfill

$$c_t = \lambda V_t - \mu(V_t). \tag{3.14}$$

The question is when we can express the value process as a function of the cash flow process. For this to be done we have to be able to invert the function $f(x) := \lambda x - \mu(x)$. A sufficient condition for the function f to be invertible is that either $\mu'(x) > \lambda$ or $\mu'(x) < \lambda$. The first case, however, will not be sufficient for our purposes, as is seen from the following. Take $\alpha > 0$ and let $\mu(x) = \alpha x$ and $\sigma(x) = \sqrt{\alpha}x$. Further let λ be such that $0 < \lambda < \alpha/2$. Then $\mu'(x) - \lambda > 0$ for all $x \in \mathbb{R}$, but

$$V_t \Lambda_t = V_0 \exp\left(\left[\frac{1}{2}\alpha - \lambda\right] t + \sqrt{\alpha} B_t\right),$$

which tends to infinity a.s. as $t \to \infty$, contradicting the fact that is should converge to 0 a.s.

Proposition 3.13 *Let $\mu \in C^2(\mathbb{R};\mathbb{R})$ and let $\sigma : \mathbb{R} \to (0,\infty)$ be such that*

$$dV_t = \mu(V_t)dt + \sigma(V_t)dB_t; \quad V_0 = v \tag{3.15}$$

possesses a strong solution. If there exists a constant λ such that $\mu'(x) < \lambda$, then there exists a unique cash flow process generating the value process via $V_t = E\left[\int_t^\infty c_s e^{-\lambda(t-s)} ds \,\middle|\, \mathcal{F}_t\right]$. Further, letting $I(x)$ denote the inverse function of $\lambda x - \mu(x)$, the cash flow process generating V has the dynamics

$$dc_t = \left[(\lambda - \mu'(I(c_t)) \cdot \mu(I(c_t)) - \frac{1}{2}\mu''(I(c_t))\sigma^2(I(c_t))\right] dt \tag{3.16}$$

$$+(\lambda - \mu'(I(c_t))) \cdot \sigma(I(c_t))dB_t. \tag{3.17}$$

Proof. We know from the earlier discussion that $c_t = \lambda V_t - \mu(V_t)$. Introduce again $f(x) = \lambda x - \mu(x)$. Since $f'(x) = \lambda - \mu'(x)$ it follows from the assumptions that $f'(x) > 0$ for every $x \in \mathbb{R}$, and hence f is invertible. The expression for the dynamics of the cash flow process follows immediately from an application of Itô's formula. □

Example 3.14 *Assume that we want the value process to have a linear drift term: $dV_t = (a + bV_t)dt + \sigma(V_t)dB_t$. We let σ be unspecified so far. Using Proposition*

3.13 we see that the cash flow process producing this drift must be $c_t = (\lambda - b)V_t - a$. *Since the derivative of the drift term is* b, *we see that if we let* λ *be any constant discount rate strictly greater than* b, *then the conditions of the theorem are fulfilled. The dynamics of* c *becomes*

$$dc_t = \{a(\lambda - b) + b(a + c_t)\}\, dt + (\lambda - b)\sigma \left(\frac{a + c_t}{\lambda - b} \right) dB_t.$$

□

Two of the most commonly used diffusions are the geometric Brownian motion and the mean reverting Ornstein-Uhlenbeck process.[1] Together with having a lognormal distribution and being strictly positive the geometric Brownian motion also has very nice computational properties. The Ornstein-Uhlenbeck process is the only stationary Gaussian process.

Proposition 3.15 *Let* $\lambda > 0$ *be a constant discount rate. If the derivative of the drift term (which in both cases below is constant) is strictly less than* λ, *then the following holds:*

(a) A value process is a geometric Brownian motion if and only if the cash flow process generating it is a geometric Brownian motion.

(b) A value process is an Ornstein-Uhlenbeck process if and only if the cash flow process generating it is an Ornstein-Uhlenbeck process.

Proof. For part (a) assume that c has dynamics $dc_t = \alpha c_t dt + \sigma c_t dB_t$, where $\alpha < \lambda$ and $\sigma > 0$. Then $V_t = c_t/(\lambda - \alpha)$ and $dV_t = \alpha V_t dt + \sigma V_t dB_t$. For the other direction we assume that V_t has dynamics given by $dV_t = \alpha V_t dt + \sigma V_t dB_t$, where again $\alpha < \lambda$ and $\sigma > 0$. The drift condition implies that $c_t = (\lambda - \alpha)V_t$, and we are finished with part (a). For (b) let a, b and σ be strictly positive real numbers and assume that the cash flow process solves the Ornstein-Uhlenbeck SDE $dc_t = a(b - c_t)dt + \sigma dB_t$. It is well known that the solution to this equation can be written

$$c_t = c_s e^{-a(t-s)} + b\left(1 - e^{-a(t-s)}\right) + \sigma \int_s^t e^{-a(t-u)} dB_u.$$

Since

$$E\left[c_s | \mathcal{F}_t\right] = b + e^{-a(t-s)}(c_s - b)$$

for $0 \le t \le t$ we have

$$V_t = E\left[\int_t^\infty e^{-\lambda(s-t)} c_s ds \middle| \mathcal{F}_t \right] = \frac{b}{\lambda} + \frac{c_t - b}{\lambda - a}.$$

[1] When we say that X is an Ornstein-Uhlenbeck process we mean that X satisfies the SDE $dX_t = a(b - X_t)dt + \sigma dB_t$ for some constants a, b and $\sigma > 0$. Strictly speaking, X is only an Ornstein-Uhlenbeck process if it has $b = 0$.

and from this

$$dV_t = a\left(\frac{b}{\lambda} - V_t\right)dt + \frac{\sigma}{\lambda - a}dB_t.$$

Now assume that we want V to be an Ornstein-Uhlenbeck process; specifically assume that the drift of V is given by $a(b - V_t)$. The cash flow process has to fulfill $c_t = \lambda V_t - ab + aV_t$, implying that

$$dc_t = a(b\lambda - c_t)dt + \sigma(\lambda + a)dB_t.$$

□

One consequence of the fact that the cash flows must follow a geometric Brownian motion if we want the value process to do so, is that if we want to model the stock price as a geometric Brownian motion and we believe that a discounted cash flow model gives the value of the stock, then the cash flows must also follow a geometric Brownian motion. Thus, the cash flows of the firm must be strictly positive, a fact that is not reasonable to assume for all firms. On the other hand we could argue that the value of the firm should be the discounted value of the dividends, and since dividends are always non-negative, we could (in theory) model them as a geometric Brownian motion.

3.4.3 Applications to real options

We end this section on Brownian models with some examples on how the methods described earlier can be specifically applied to problems arising in the valuation of real options. The idea of real options is that added to the net present value (represented by the value process as specified here) there should be a value coming from some implicit option. A typical example is the case when we own a gold mine. Suppose that the gold price is so low that it is not profitable (in the sense that the value process at this instant is negative) to keep the mine running. There is, however, a possibility that the gold price will increase in the future, and it is possible that it eventually will become profitable to use the mine. Thus, we can see the mine as an option with the gold price as the underlying asset, and as other options it has a value even though it is not presently in the money. The value added to the mine in this case is the value of waiting to invest. If we have a 'now-or-never' choice today to decide if should close the mine down or let it run, we should of course (still assuming that the value process today is negative) shut the mine down. Examples where there exists an embedded real option are many, ranging from investment timing (when an irreversible investment should be done), entering and exiting markets, sequential investments (often an investment is done in stages with a possibility of interrupting after the first stage if it is no longer profitable, the search for a new drug at a pharmaceutical company is a typical example) and real estate (where unexploited land can increase in value if the rent increase and/or the cost of construction decrease) to purely noneconomic applications such as legal reforms (since there is a cost present, both monetary and socially, when some laws are changed, there is a value in waiting to see what the opinion among the voters is). These examples (see Dixit & Pindyck [9]) serve to show that the area to which we may apply real options is indeed vast.

There are two main routes to follow when modelling the value of a project. Either we directly model the value process, or we model the cash flow process generating the value process and then derive the properties of the value process from the cash flows.

Example 3.16 *When modelling the value process directly, the geometric Brownian motion is often used. As was pointed out above it guarantees among other things that the value process is strictly positive. To get a model with mean-reverting value process, the following models are sometimes used (b and η are positive constants)*

- $dV_t = \eta(b - V_t)V_t dt + \sigma V_t dB_t$,
- $dV_t = \eta(b - \ln V_t)V_t dt + \sigma V_t dB_t$, *and*
- $dV_t = \eta(b - V_t)dt + \sigma V_t dB_t$.

The second equation describes the dynamics of an exponential Ornstein-Uhlenbeck process; see Sick [18] for these models. In all these cases, if we start at a strictly positive value at time 0 we will have $V_t \geq 0$ a.s. for every $t \geq 0$. □

One way to guarantee a positive value process is to have a positive cash flow process. It is, however, not a necessary condition, as we will now see. Assume that the discount rate is a positive constant λ, and that we want V to be strictly positive and have dynamics given by $dV_t = a(b - V_t)dt + \sigma(V_t)dB_t$, were we wait to specify σ. Here a and b are positive constants. We thus want to solve $\lambda V_t - c_t = a(b - V_t)$, yielding $c_t = (\lambda + a)V_t - ab$. From this do we get

$$dc_t = a(b\lambda - c_t)dt + (\lambda + a)\sigma\left(\frac{c_t + ab}{\lambda + a}\right) dB_t.$$

In general c can take both positive and negative values. Assuming that we still want V to be strictly positive, take a continuous function $g : [0, \infty) \to [0, \infty)$ such that $g(0) = 0$ and let $\sigma(V_t) = g(V_t)$. As an explicit example we can take $g(x) = x^\gamma$, $\gamma > 0$, yielding

$$dc_t = a(b\lambda - c_t)dt + (\lambda + a)^{1-\gamma}(c_t + ab)^\gamma dB_t.$$

In some cases it could be desirable to allow for the value process to take negative values. Either we let the value process continue, disregarding simply whether it is positive or negative, or we could start the cash flow process, letting it generate the value process, but consider the project bankrupt at the first time τ with $V_\tau \leq 0$.

Another application to real options of the relation between the cash flow process and the value process concerns estimation of the volatility. Assuming that a cash flow process is driving the value process (and not assuming the value process itself as the underlying object) we have to be able to estimate the diffusion term of c. The problem is that the cash flow process is often not directly observable; what we observe is the value process. As an example we could think of a pharmaceutical company whose survival is dependent on the success of a new drug. Research on this drug is still done, and it is not certain whether it will be good enough or not. We could try to value this company using real options, in which case the dynamics of the cash flow process is needed. But what we observe, as was said earlier, is the value process. We can assume that the stock price of this company is equal to value process, use the time series of the stock price to estimate parameters for V and then, using the relation $\mu(V_t) - \lambda V_t = c_t$, estimate the parameters for c.

Bibliography

[1] Armerin, F. (2002), 'Valuation of Cash Flows in Discrete Time', Working paper

[2] Bias, B., Björk, T., Cvitanić, J., El Karoui, N., Jouini, E.& Rochet, J. C. (1997), 'Financial Mathematics', Lecture Notes in Mathematics, 1656, *Springer-Verlag*

[3] Björk, T. (1998), 'Arbitrage Theory in Continuous Time', *Oxford University Press*

[4] Brealey, R. A. & Myers, S. C. (1991), 'Principles of Corporate Finance', Fourth Edition, *McGraw-Hill*

[5] Cochran, J. H. (2001), 'Asset Pricing', *Princeton University Press*

[6] Copeland, T. & Antikarov, V (2001), 'Real Options', *Texere*

[7] Copeland, T., Koller T. & Murrin, J. (2000), 'Valuation', *John Wiley & Sons, Inc.*

[8] Delbaen, F. & Schachermayer, W. (1994), 'A General Version of the Fundamental Theorem of Asset Pricing', *Math. Annalen*, Vol. 300, 463-520

[9] Dixit, A. K. & Pindyck, R. S. (1994), 'Investment under Uncertainty', *Princeton University Press*

[10] Duffie, D. (1996), 'Dynamic Asset Pricing Theory', *Princeton University Press*

[11] Duffie, D., Ma, J. & Yong, J. (1995), 'Black's consol rate conjecture', *The Annals of Applied Probability*, Vol. 5, No. 2, 356-382

[12] Durrett, R. (1996), 'Probability: Theory and Examples, Second edition, *Duxbury Press*

[13] El Karoui, N., Peng S. and Quenez, M. C. (1997), 'Backward stochastic differential equations in finance', *Mathematical Finance* 7, No. 1, 1-71

[14] Elliott, R. J. (1982), 'Stochastic Calculus and Applications', *Springer-Verlag*

[15] Harrison, J. M. & Kreps, M. D. (1979), 'Martingales and Arbitrage in Multiperiod Securities Markets', *Journal of Economic Theory*

[16] Hubalek F. & Schachermayer, W. (2001), 'The Limitations of No-Arbitrage Arguments for Real Options', *International Journal of Theoretical and Applied Finance* Vol. 4, No. 2

[17] Jacod, J. & Shiryaev, A. N. (2003), 'Limit Theorems for Stochastic Processes' (2nd Ed.), *Springer-Verlag*

[18] Jarrow, R. A., Maksimovic, V. & Ziemba, W. T. (Editors) (1995), 'Finance', Handbooks in Operations Research and Management Science, *North-Holland*

[19] Lépingle, P. (1978), 'Sur le compartement asymptotique des martingales locales', *Springer Lecture Notes in Mathematics* 649, 148-161

[20] Ma, J. & Yong, J. (1999), 'Forward-Backward Stochastic Differential Equations and Their Applications', *Springer-Verlag*

[21] Martin-Löf, A. (1986), 'A Theory of Life Insurance', *Scandinavian Actuarial Journal*, No. 2, p. 65-81

[22] Norberg, R. (1990), 'Payment Measures, Interest, and Discounting – An Axiomatic Approach with Applications to Insurance *Scandinavian Actuarial Journal* p. 14-33

[23] Norberg, R. (1991), 'Reserves in Life and Pension Insurance', *Scandinavian Actuarial Journal* p. 3-24

[24] Norberg, R. (2001), 'Financial Mathematics in Life and Pension Insurance', *Lecture notes*

[25] Protter, P. (2004), 'Stochastic Integration and Differential Equations', Second Edition, *Springer-Verlag*

[26] Rogers, L. C. G. (1997), 'The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates', *Mathematical Finance* 7, 157-176

[27] Øksendal, B. (1998), 'Stochastic Differential Equations', Fifth Edition, *Springer-Verlag*

Chapter 4

Immunization of Deterministic and Stochastic General Cash Flows

4.1 Introduction

A common situation in immunization theory is to consider the case where the cash flows of both the assets and liabilities occur at a finite set of times. To illustrate the methodology, consider a portfolio consisting of assets whose non-negative cash flows $(A_j)_{1\leq j\leq m}$ occur at dates $0 \leq t_1 < t_2 < \ldots < t_m < \infty$, and liabilities whose non-negative cash flows $(L_k)_{1\leq k\leq n}$ occur at times $0 \leq s_1 < s_2 < \ldots < s_n < \infty$. This is a deterministic model, so all cash flows are known at time 0. We assume that the discount factor from time $s \geq 0$ back to time 0 is given by $v(s)$. The present value of the assets equals the present value of the liabilities if and only if

$$\sum_{j=1}^{m} A_j v(t_j) = \sum_{k=1}^{n} L_k v(s_k), \tag{4.1}$$

and we assume that we have chosen the asset cash flows such that this condition is fulfilled. The immunization problem concerns the present value of the assets and the liabilities under a new term structure of interest rates given by discount factors v'. We say that the portfolio is immunized against the new term structure v' if

$$\sum_{j=1}^{m} A_j v'(t_j) \geq \sum_{k=1}^{n} L_k v'(s_k).$$

If we consider a single liability (i.e. there is one time s_1 at which the only cash flow generated by the liabilities occur), then the portfolio is immunized against every new term structure v' such that $f(s) = v'(s)/v(s)$, known as the shift function, is convex, if we have chosen the cash flows of the assets such that

$$\sum_{j=1}^{m} t_j A_j v(t_j) = s_1 L_1 v(s_1)$$

(duration matching) in addition to condition (4.1). This follows immediately from Jensen's inequality; see e.g. Montrucchio & Peccati [5]. When we face multiple liabilities, then we are immunized against any convex shift (i.e. against every convex shift function) if and only if we can decompose the asset portfolio into n sub-portfolios, where each of these sub-portfolios immunizes one liability cash flow L_k, $1 \leq l \leq n$ each. Such a decomposition is known as a Shiu decomposition (Uberti [9]).

Going back to the case with multiple liabilities, let

$$\mu_A(t_j) = \frac{A_j v(t_j)}{\sum_{j=1}^{m} A_j v(t_j)}, \ j = 1, 2, \dots m,$$

and

$$\mu_L(s_k) = \frac{L_k v(s_k)}{\sum_{k=1}^{n} L_k v(s_k)}, \ k = 1, 2, \dots n.$$

Then μ_A and μ_L defines two probability measures on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. We will refer to these measures as the asset and liability measure respectively. If we denote by X_A and X_L two random variables with distribution μ_A and μ_L respectively, then we can write the duration D_A of the cash flows of the assets and the duration of the cash flows of the liabilities D_L as

$$D_A = E[X_A] \text{ and } D_L = E[X_L]$$

respectively. The duration matching condition can thus be recast as

$$E[X_A] = E[X_L].$$

In the case of a single liability the duration is simply the time, say s_1, at which the liability is due. Now assume that the discount factor is given by $v(s) = e^{-\delta s}$ for some $\delta > 0$. The present value of the liabilities is then

$$V_L(\delta) = \sum_{k=1}^{n} L_k e^{-\delta s_k}.$$

The derivative of V_L with respect to δ is

$$\frac{dV_L}{d\delta} = -V_L \cdot D_L,$$

and this implies

$$V_L(\delta + \varepsilon) \approx V_L(\delta)(1 - D_L(\delta)).$$

It is straightforward to generalize the previous reasoning to asset and liability measures that not necessarily have support on some finite set of times, but more generally are probability measures on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ with finite support. If we want to consider asset and liability measures with infinite support (i.e. the whole positive real line), care has to be taken so that the value of the cash flow streams are finite. Although the present value of e.g. the assets under the new term structure may be finite, we are not guaranteed that the present value of the liabilities under the new term structure is finite, and vice versa. To circumvent this problem, we can either consider only assets and liabilities with finite support, or restrict the class of new term structures. One should note that by choosing $A = L$ we are immunized against any shift, but this choice is not always possible or desirable.

It is not unusual for the cash flows of the assets and/or the liabilities to be random. We will also consider this case. A stream L of random liability cash flows are modelled by an increasing process. The present value of L under the set of discount factors v is defined as

$$E\left[\int_{[0,\infty)} v(s)dL_s\right].$$

Here v is a deterministic discount factor.

There has sometimes been argued that an immunized position is an arbitrage opportunity, a view we disagree with. By looking at immunization as a static approach, in which the change of discount factor is made at time 0 and not due to the fact that time passes, we have no problems with possible arbitrage opportunities. We should regard immunization theory as a sort of stress testing, testing what will happen with the present values of the cash flows if another set of discount factors is used.

The rest of this paper is organized as follows. We will first present immunization results for the case with deterministic cash flows. The main interest will be on immunization of a single liability. We show that old immunization results carry over to our general case, the main difference being that we must take care so that the value under the new term structure is finite. Some new lower bounds on the difference between the present values of the assets and the liabilities will also be presented. Then we will consider the case with stochastic cash flows. Here the focus will also be on single liability immunization. It turns out that it is not possible to generalize deterministic immunization results to the stochastic case. One exception is when a random amount is to be paid out on a deterministic date. In this case the deterministic immunization results hold.

4.2 Deterministic immunization

The classical literature on immunization consider the situation when the cash flows of both the assets and liabilities occur at a finite set of points in time. It is straightforward to generalize this to asset and liability measures with finite support. If we want to consider asset and liability measures with infinite support, care has to be taken so that the value of the cash flow streams are finite. In fact, although the present value of the assets under the new term structure may be finite, we are not guaranteed that the present value of the liabilities under the new term structure is finite, and vice versa. To circumvent this problem, we need to restrict the class of new term structures.

Definition 4.1 *A* discount function *is a (Borel) measurable and locally bounded function* $v : [0, \infty) \to (0, \infty)$.

Since every discount function v is strictly positive $R = -\ln v$ is well defined. Well defined is also

$$Y(s) = \begin{cases} R(s)/s & \text{when } s > 0 \\ \xi & \text{when } s = 0, \end{cases}$$

where $\xi \in \mathbb{R}$. Y is the yield curve at time 0, i.e. $Y(s)$ is the yield-to-maturity at time 0 of a non-defaultable zero coupon bond with maturity s. We can write

$$v(s) = e^{-R(s)} = e^{-Y(s)s} \quad \text{for } s \geq 0.$$

Since $\xi \cdot s = 0$, the numerical value of ξ does not matter, and the value ξ is 'at our disposal'. If $R(s)/s$ converges as $s \downarrow 0$, then it is reasonable to let ξ be this limit. As an example, assume that we have the representation

$$v(s) = e^{-\int_0^s i(u)du} \quad \text{for } s \geq 0$$

for some continuous function $i : \mathbb{R}_+ \to \mathbb{R}$. The function i represents the forward rates at time 0. In this case we have

$$\frac{R(s)}{s} = \frac{1}{s}\int_0^s i(u)du \to i(0) \quad \text{as } s \downarrow 0.$$

We model the asset and liability cash flows as two increasing functions:

Definition 4.2 *By an* increasing function *we mean a finite non-decreasing function* $g : [0, \infty) \to [0, \infty)$ *that is right-continuous and null at the origin (i.e.* $g(0-) = g(0) = 0$*).*

Every increasing function g induces a σ-finite measure μ_g on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ by

$$\mu_g[0, t] = g(t).$$

We write $dg = d\mu_g$. Since every increasing function $g : \mathbb{R}_+ \to \mathbb{R}$ has at most a countable number of jumps on each compact set we can decompose g as

$$g(t) = \sum_{0 \leq s \leq t} \Delta g(s) + \left[g(t) - \sum_{0 \leq s \leq t} \Delta g(s) \right] =: \sum_{0 \leq s \leq t} \Delta g(s) + g_c(t).$$

Here g_c is the continuous part of g. Using the Lebesgue decomposition we can write g_c as the sum of a singular and an absolute continuous part:

$$g_c(t) = g_s(t) + \int_0^t \dot{g}_c(s)ds.$$

A jump in the increasing process is due to a lump-sum payment, and the absolutely continuous part represents a continuous inflow of money.

Definition 4.3 *Given an increasing function g and a discount function v, the* value V *of g is given by*

$$V = \int_{[0,\infty)} v(s) dg(s).$$

If the value of g is finite we define the duration D *of g by*

$$D = \frac{\int_{[0,\infty)} s v(s) dg(s)}{\int_{[0,\infty)} v(s) dg(s)} = \int_{[0,\infty)} s \frac{v(s)}{V} dg(s).$$

Since a discount function is measurable and positive V is well defined and fulfills $0 \leq V \leq \infty$. To shorten the notation we will write $g_s = g(s)$ for the increasing functions we use as integrators.

The cash flows generated by the assets and the liabilities will be denoted A and L respectively. We think of the liabilities as given to us, while we can choose the cash flows generated by the assets as we wish. Of course the choice $A_s = L_s$ for every $s \geq 0$ will make the difference between discounted cash flows of the assets and liabilities to always be zero, regardless of term structure. It is, however, not always possible or desirable to create such perfect cash flow matching. The program now is the following:

1. Given to us is a strictly positive and locally bounded measurable function v defined on $\mathbb{R}_+$ and representing the initial term structure.

2. Given is also an increasing function L representing the cash flows of the liabilities. This L is assumed to fulfill

$$\int_{[0,\infty)} v(s) dL_s < \infty.$$

(If this was not true, we would face liabilities of infinite value, a position we would most likely reject.)

3. We finance the cash flows of the liabilities by choosing an increasing function A representing the cash flows of the assets such that

$$\int_{[0,\infty)} v(s) dA_s = \int_{[0,\infty)} v(s) dL_s.$$

4. Assume that the term structure changes to v', where v' is another discount function, and fulfills

$$\int_{[0,\infty)} v'(s) dA_s < \infty \quad \text{and} \quad \int_{[0,\infty)} v'(s) dL_s < \infty. \tag{4.2}$$

Remark 4.4 By v' we do not mean the derivative of v, but the new term structure. This notation is not optimal, but is in conformity with the one used in previous work on immunization.

The steps described above are the same as when we have asset and liability measures with finite support, except that, in this general case, we have to guarantee that the present value of the assets and liabilities are finite under the new term structure. As a short hand notation we introduce

$$V_A = \int_{[0,\infty)} v(s)dA_s \text{ and } V_L = \int_{[0,\infty)} v(s)dL_s$$

as well as

$$V'_A = \int_{[0,\infty)} v'(s)dA_s \text{ and } V'_L = \int_{[0,\infty)} v'(s)dL_s.$$

We call the following the *classical immunization problem.*

> Given that
>
> $$V_A = V_L,$$
>
> under what conditions will
>
> $$V'_A - V'_L = \int_{[0,\infty)} v'(s)dA_s - \int_{[0,\infty)} v'(s)dL_s$$
>
> be non-negative and give precise conditions on this change of value if it is negative. Here $v'(s)$ is a new set of discount functions and we assume that both V'_A and V'_L are finite.

The crucial point to note here is that we do not let time pass, but consider the value of the streams of cash flows under another discount function at time zero. Therefore classical immunization should be viewed only as a robustness test. A portfolio fulfilling

$$\int_{[0,\infty)} v'(s)dA_s \geq \int_{[0,\infty)} v'(s)dL_s$$

for every v' such that both integrals are finite will be referred to as an *immunized position.* Letting $f(s) = v'(s)/v(s)$ denote the *shift function* we see that we can write the difference between the value of the assets and the value of the liabilities with the new discount factors (under the assumption that the expectations below exist) as

$$\begin{aligned} V'_A - V'_L &= V_A \int_{[0,\infty)} f(s)\frac{v(s)}{V_A}dA_s - V_L \int_{[0,\infty)} f(s)\frac{v(s)}{V_L}dL_s \\ &= V_A \int_{[0,\infty)} f(s)d\mu_A(s) - V_L \int_{[0,\infty)} f(s)d\mu_L(s) \\ &= V_A E\left[f(X_A)\right] - V_L E\left[f(X_L)\right]. \end{aligned}$$

Here

$$\mu_A(s) = \frac{v(s)}{V_A}dA_s \text{ and } \mu_L(s) = \frac{v(s)}{V_L}dL_s$$

are two probability measures on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ and X_A and X_L are two random variables having distribution μ_A and μ_L respectively. Under the assumption that $V_A = V_L$ we see that the immunization problem is a problem concerning the properties of $E[f(X_A) - f(X_L)]$ for f in some given class of functions.

The question of which shift functions f that are admissible, in the sense that the value of assets and liabilities are finite under the new term structure defined by the shift function, is of course crucial. It is easy to see that every measurable and (uniformly) bounded function f will work. Important is whether polynomials of a given degree will work, since this is connected to the classical measures of cash flows such as duration, convexity and M-square index. We will have to assume that polynomial shift functions of sufficiently high degree are admissible. Of course, if the support of the cash flows is finite, then every shift function will be admissible. Since $f = v'/v$ we have

$$f(s) = e^{-(R'(s)-R(s))} = e^{-(Y'(s)-Y(s))s} \quad \text{for } s \geq 0.$$

By letting $\Delta(s) = Y'(s) - Y(s)$ we have

$$f(s) = e^{-\Delta(s)s} \quad \Leftrightarrow \quad \Delta(s) = -\frac{1}{s}\ln f(s) \quad \text{for} \quad s \geq 0.$$

The interpretation of Δ is that if the shift function is f, then the associated Δ is the shift of the yield curve.

4.2.1 Single liability immunization

An important (both from a mathematical and practical point of view) case is when the liability is given by

$$L_s = K\mathbf{1}(s \geq T), \quad s \geq 0,$$

where $K > 0$ is a given constant and $T > 0$ is the deterministic time at which the liability will occur. The value of this liability at 0 is

$$V_L = \int_{[0,\infty)} v(s)dL_s = Kv(T).$$

The associated liability measure is a point mass at $s = T$. Thus

$$E[f(X_L)] = f(T) = f(E[X_L]).$$

To simplify the notation we will write

$$F(v, A) = \frac{1}{v(T)} \int_{[0,\infty)} v(s)dA_s$$

for the future value of the assets with cash flow stream A at time T when the discount function is v. Expressed in terms of F a position is immunized if $F(v', A) \geq F(v, A)$. Since $F(v, A) > 0$ we can in this case with a single liability rephrase the immunization problem as follows:

Given that

$$F(v, A) = K,$$

under what conditions will

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq 0$$

be non-negative and give precise conditions on this change of value if it is negative. Here v' is a new set of discount functions and we assume that $F(v', A)$ is finite.

First of all we will show a general result giving a lower bound for the change in value. It is a generalization of a result by Fong & Vasicek [1] (recall that $f(s) = e^{-\Delta(s)s}$).

Proposition 4.5 *Let v be a discount function, and let L be a single liability due at $T > 0$ with amount K. If A is a cash flow process such that $F(v, A) = K$, then*

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq -\big(E\left[\Delta(X_A)X_A\right] - \Delta(T)T\big)$$

for every v' such that $F(v', A) < \infty$.

Proof. Using the inequality $e^x \geq 1 + x$, which is true for every $x \in \mathbb{R}$, we get

$$\begin{aligned} F(v', A) &= \frac{1}{v'(T)} \int_{[0,\infty)} v'(s) dA(s) \\ &= \frac{1}{v(T)} \int_{[0,\infty)} \frac{v(T)}{v'(T)} \frac{v'(s)}{v(s)} v(s) dA(s) \\ &= \frac{1}{v(T)} \int_{[0,\infty)} e^{\Delta(T)T - \Delta(s)s} v(s) dA(s) \\ &\geq \frac{1}{v(T)} \int_{[0,\infty)} \{1 + \Delta(T)T - \Delta(s)s\}\, v(s) dA(s) \\ &= F(v, A)\left(1 - E\left[\Delta(X_A)X_A - \Delta(T)T\right]\right), \end{aligned}$$

We have used the fact that we have value matching: $F(v, A) = V_A = V_L$. Rearranging yields

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq -\big(E\left[\Delta(X_A)X_A - \Delta(T)T\right]\big).$$

□

The following corollary is an immediate consequence of the previous proposition.

Corollary 4.6 *Assuming the same conditions as in the previous proposition, any duration matched position is immunized if $f(s)$ is convex.*

Proof. The fact that e^x is increasing and convex implies that f is convex if and only if $\Delta(s)s$ is concave. Using that $D_A = E[X_A] = T$ when the position is duration matched and the convexity of $f(s)$ we get

$$\begin{aligned}\frac{F(v', A) - F(v, A)}{F(v, A)} &\geq E[-\Delta(X_A)X_A] + \Delta(T)T \\ &\geq -\Delta(E[X_A])E[X_A] + \Delta(T)T = 0\end{aligned}$$

from Jensen's inequality. □

Thus, the previous corollary shows that a duration matched position of a single liability is immunized against every admissible convex shift function. We can write the lower bound in Proposition 4.5 as

$$-\Delta(T)E[X_A - T] - E[X_A(\Delta(X_A) - \Delta(T)].$$

The first term in this expression is an 'asset component' (it is small if $EX_A \approx T$) and the second term is a 'yield component' (it is small if the change in yield ($= Y' - Y = \Delta$) is small). If Δ is differentiable we have the following heuristic analysis. It follows from

$$\Delta(X_A) \approx \Delta(T) + \Delta'(T)(X_A - T)$$

that we have the approximate lower bound

$$-\Delta(T)E[X_A - T] - \Delta'(T)E\left[X_A^2 - TX_A\right].$$

If we are duration matched then $E[X_A] = T$ and we have the approximate lower bound

$$-\Delta'(T)\text{Var}(X_A).$$

We can conclude that in the in the case when Δ is differentiable, the lower bound is (approximately) determined by the derivative of Δ at T and the variance of the asset measure. There is, however, no reason to exclude Δ from being very irregular, especially we do not always assume Δ to be even continuous.

Corollary 4.7 *Assume the same conditions as in Proposition 4.5. If the change in yield Δ is uniformly bounded by M and the position is duration matched, then*

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq -2TM.$$

If additionally Δ is Lipschitz continuous with constant C, then

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq -(M + TC)E|X_A - T|.$$

Proof. We have

$$\begin{aligned}\frac{F(v',A)-F(v,A)}{F(v,A)} &\geq -E\left[\Delta(X_A)X_A\right]+\Delta(T)T \\ &= -\Delta(T)E\left[X_A-T\right]-E\left[X_A(\Delta(X_A)-\Delta(T))\right] \\ &\geq -E\left[X_A|\Delta(X_A)-\Delta(T)|\right] \\ &\geq -2MT.\end{aligned}$$

To show the other inequality we use the fact that the assumption of Lipschitz continuity implies $\Delta(X_A)-\Delta(T)\geq -C|X_A-T|$:

$$\begin{aligned}\frac{F(v',A)-F(v,A)}{F(v,A)} &\geq -E\left[\Delta(X_A)X_A\right]+\Delta(T)T \\ &= -E\left[\Delta(X_A)(X_A-T)\right]-TE\left[(\Delta(X_A)-\Delta(T))\right] \\ &\geq -ME|X_A-T|-TC\,E|X_A-T| \\ &\geq -(M+TC)E|X_A-T|.\end{aligned}$$

□

Since $E[X_A]=T$ under duration matching, the factor $E|X_A-T|$ in the last inequality is the central absolute deviation of X_A.

We will now consider the case when the discount function can be written $v(s)=\exp\left(-\int_0^s i(u)du\right)$ for some locally bounded measurable function i. With $\varepsilon=i'-i$ the shift function now takes the form

$$f(s)=e^{-\int_0^s(i'(u)-i(u))du}=e^{-\int_0^s\varepsilon(u)du}\quad\text{for } s\geq 0.$$

Expressing the lower bound in terms of ε instead of Δ means that we look at changes in forward rates instead of changes in yield.

Proposition 4.8 *Let v be given by $v(s)=\exp\left(-\int_0^s i(u)du\right)$. Then v is a discount function. With $v'(s)=\exp\left(-\int_0^s i'(u)du\right)$ for some new forward rates i' and with $\varepsilon=i'-i$, we have*

$$\frac{F(v',A)-F(v,A)}{F(v,A)}\geq -E\left[\int_T^{X_A}\varepsilon(u)du\right]. \tag{4.3}$$

Proof. Obviously v defined as above will be locally bounded and measurable, so v is a discount function. We have

$$\Delta(s)s=-\log f(s)=\int_0^s\varepsilon(u)du,$$

and the lower bound given by Proposition 4.5 becomes

$$\begin{aligned}\frac{F(v',A)-F(v,A)}{F(v,A)} &\geq -E\left[\int_0^{X_A}\varepsilon(s)ds-\int_0^T\varepsilon(s)ds\right]\\ &= -E\left[\int_T^{X_A}\varepsilon(s)ds\right].\end{aligned}$$

□

If

$$\int_\Omega\int_T^{X_A(\omega)}|\varepsilon(u)|du\,dP(\omega)\equiv\int_{[0,\infty)}\int_T^x|\varepsilon(u)|du\,d\mu_A(x)<\infty$$

then we can change the order of integration to arrive at

$$-E\left[\int_T^{X_A}\varepsilon(u)du\right]=-\int_0^\infty\varepsilon(u)\big[P(X_A>u)-\mathbf{1}(u\leq T)\big]du.$$

This shows that given a shift ε of the forward rates, the lower bound on the difference between the present value of the assets and the liabilities depends on where on $[0,\infty)$ ε puts weight. Especially note that when we have a constant shift the lower bound is zero. By making additional assumptions on the class of functions to which ε belongs we can get more information out of this lower bound. The following corollary gives some examples. They are all consequences of Equation (4.3). The proof of number (3) is a generalization of a result in Fong & Vasicek [1].

Corollary 4.9 *Given that ε is such that $F(v',A)$ is finite and the position is duration matched, the following holds:*

(1) If $\int_0^s\varepsilon(u)du$ is concave on $[0,\infty)$, then

$$\frac{F(v',A)-F(v,A)}{F(v,A)}\geq 0.$$

(2) If $\varepsilon(s)$ is uniformly bounded by M then

$$\frac{F(v',A)-F(v,A)}{F(v,A)}\geq -ME|X_A-T|.$$

(3) If $d\varepsilon/ds$ exists for every $s\geq 0$ and is bounded from above by k, then

$$\frac{F(v',A)-F(v,A)}{F(v,A)}\geq -\frac{kE\left[(X_A-T)^2\right]}{2}.$$

(4) If $\varepsilon(s) = \sum_{\ell=0}^{n} a_\ell s^\ell$*, then*

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq -\sum_{\ell=0}^{n} \frac{a_\ell}{\ell + 1} \left(E\left[X_A^{\ell+1}\right] - T^{\ell+1} \right).$$

(5) Let $p : \mathbb{R} \to \mathbb{R}$ *be an integrable function. If* $|\varepsilon(s) - p(s)| \leq M$ *for every* $s \geq 0$ *and some* $M \geq 0$ *then*

$$\frac{F(v', A) - F(v, A)}{F(v, A)} \geq -E\left[M|X_A - T| + \int_T^{X_A} p(u)du\right].$$

Proof.

(1) This follows from Corollary 4.6.

(2) First write

$$-E\left[\int_T^{X_A} \varepsilon(u)du\right] = -E\left[\mathbf{1}(T \leq X_A)\int_T^{X_A} \varepsilon(u)du\right] - E\left[\mathbf{1}(X_A < T)\int_T^{X_A} \varepsilon(u)du\right].$$

We have the inequalities

$$E\left[\mathbf{1}(T \leq X_A)\int_T^{X_A} \varepsilon(u)du\right] \leq ME\left[\mathbf{1}(T \leq X_A)(X_A - T)\right]$$

and

$$\begin{aligned} E\left[\mathbf{1}(X_A < T)\int_T^{X_A} \varepsilon(u)du\right] &= E\left[\mathbf{1}(X_A < T)\int_{X_A}^{T} (-\varepsilon(u))du\right] \\ &\leq ME\left[\mathbf{1}(X_A < T)(T - X_A)\right]. \end{aligned}$$

We get

$$\begin{aligned} -E\left[\int_T^{X_A} \varepsilon(u)du\right] &\geq -ME\left[\mathbf{1}(T < X_A)|X_A - T| + \mathbf{1}(X_A \leq T)|T - X_A|\right] \\ &= -ME|X_A - T|. \end{aligned}$$

(3) Using partial integration we can write

$$-\int_T^{X_A} \varepsilon(u)du = \varepsilon(T)(T - X_A) - \int_{X_A}^{T} (u - X_A)\varepsilon'(u)du.$$

Now

$$\int_{X_A}^{T} (u - X_A)\varepsilon'(u)du \leq k\frac{(X_A - T)^2}{2} \text{ on } \{X_A < T\}$$

and

$$\int_{X_A}^{T}(u-X_A)\varepsilon'(u)du=\int_{T}^{X_A}(X_A-u)\varepsilon'(u)du\le k\frac{(X_A-T)^2}{2}\text{ on }\{T\le X_A\}.$$

Hence

$$-E\left[\int_T^{X_A}\varepsilon(u)du\right]\ge -E\left[k\frac{(X_A-T)}{2}\right].$$

(4) We get

$$-E\left[\int_T^{X_A}\varepsilon(u)du\right]=-E\left[\int_T^{X_A}\sum_{\ell=0}^{n}a_\ell u^\ell du\right]=-\sum_{\ell=0}^{n}\frac{a_\ell}{1+\ell}\left(E\left[X_A^{\ell+1}\right]-T^{\ell+1}\right).$$

(5) As in the proof of (2) we look at the cases $T\le X_A$ and $X_A<T$. We get

$$E\left[\mathbf{1}(T\le X_A)\int_T^{X_A}\varepsilon(u)du\right]\le E\left[\mathbf{1}(T\le X_A)\left\{M(X_A-T)+\int_T^{X_A}p(u)du\right\}\right]$$

and

$$E\left[\mathbf{1}(X_A<T)\int_T^{X_A}\varepsilon(u)du\right]\le E\left[\mathbf{1}(X_A<T)\left\{M(T-X_A)+\int_{X_A}^{T}(-p(u))du\right\}\right].$$

This gives

$$-E\left[\int_T^{X_A}\varepsilon(u)du\right]\ge -E\left[M|X_A-T|+\int_T^{X_A}p(u)du\right].$$

□

4.2.2 General liability immunization

We will now look at immunization when the liabilities are given by a general increasing function. Recall that if $V_A=V_L$, then

$$\begin{aligned}V_A'-V_L' &= V_A\left(E\left[f(X_A)\right]-E\left[f(X_L)\right]\right)\\ &= V_A\left(\int_{[0,\infty)}f(s)\frac{v(s)}{V_A}dA(s)-\int_{[0,\infty)}f(s)\frac{v(s)}{V_L}dL(s)\right)\end{aligned}\qquad(4.4)$$

for every shift function f such that the expectations exist. The results in this section are generalizations of results from Shiu [7], [8] and Hürlimann [3] (see also Chapter 3 in [6]).

Example 4.10 *Assume that the discount rate is constant at the level δ. Thus we have*

$$V_A(\delta) = \int_{[0,\infty)} e^{-\delta s} dA_s \quad \text{and} \quad V_L(\delta) = \int_{[0,\infty)} e^{-\delta s} dL_s.$$

Assuming A and L being so 'nice' that we can interchange differentiation and integration, the first and second order derivative of the present value of the assets with respect to δ are

$$V_A'(\delta) = -\int_{[0,\infty)} s e^{-\delta s} dA_s \quad \text{and} \quad V_A''(\delta) = \int_{[0,\infty)} s^2 e^{-\delta s} dA_s$$

respectively. We get similar expressions for V_L' and V_L''. Under value and duration matching, the difference $N(\delta) = V_A(\delta) - V_L(\delta)$ fulfills

$$N(\delta) = 0 \quad \text{and} \quad N'(\delta) = 0.$$

If furthermore the cash flows of the assets and liabilities are chosen such that

$$\int_{[0,\infty)} s^2 e^{-\delta s} dA_s \geq \int_{[0,\infty)} s^2 e^{-\delta s} dL_s,$$

then

$$N''(\delta) \geq 0.$$

This means that whatever change of the constant rate that occurs, the position is immunized. □

In view of this example, it looks like the general liability case is (from an immunization point of view) essentially like the single liability case. The problem is that the result in the example heavily relies on the flat yield curve assumption. To see this, consider the general case as in Equation (4.4) with a convex shift function f. We can then use Jensen's inequality on the first term. In the single liability case this is enough, but it will not be enough when we face general liabilities.

Shift functions that are C^2

In this section we will assume that the shift function is twice continuously differentiable. We can then write

$$f(s) = f(0) + f'(0)s + \int_0^\infty (s-x)^+ f''(x) dx$$

for $s \geq 0$. The new value of the assets is

$$\begin{aligned} V_A' &= \int_{[0,\infty)} \left\{ f(0) + f'(0)s + \int_0^\infty (s-x)^+ f''(x)dx \right\} v(s) dA_s \\ &= f(0)V_A + f'(0)V_A D_A + V_A \int_{[0,\infty)} \int_0^\infty (s-x)^+ f''(x) dx \, \frac{v(s)}{V_A} dA(s). \end{aligned}$$

An analogous relation holds for V_L'. Assume $V_A' < \infty$ and $V_L' < \infty$ and insert these expressions into equation (4.4) and assuming the same value and duration (both also assumed finite) of the assets and liabilities gives

$$\begin{aligned} V_A' - V_L' &= V_A \int_{[0,\infty)} \int_0^\infty (s-x)^+ f''(x) dx \, d(\mu_A - \mu_L)(s) \qquad (4.5) \\ &= V_A \int_0^\infty f''(x) \int_{[0,\infty)} (s-x)^+ d(\mu_A - \mu_L)(s) dx \\ &= V_A \int_0^\infty f''(x) E\left[(X_A - x)^+ - (X_L - x)^+\right] dx. \qquad (4.6) \end{aligned}$$

The Mean Value Theorem for integrals implies that we can write

$$\begin{aligned} \int_0^\infty E\left[(X_A - x)^+\right] f''(x) dx &= f''(\zeta_A) \int_0^\infty E\left[(X_A - x)^+\right] dx \\ &= f''(\zeta_A) E\left[\frac{X_A^2}{2}\right] \end{aligned}$$

for some $\zeta_A \in [0, \infty)$. The same relation holds for the liabilities, with some ζ_L instead of ζ_A, and we get

$$V_A' - V_L' = \frac{V_A}{2}\Big(f''(\zeta_A) E\left[X_A^2\right] - f''(\zeta_L) E\left[X_L^2\right]\Big). \qquad (4.7)$$

We let

$$C_A = E\left[X_A^2\right] \quad \text{and} \quad C_L = E\left[X_L^2\right]$$

denote the *convexity* of the assets and liabilities respectively.

Proposition 4.11 *If $V_A = V_L < \infty$ and $D_A = D_L < \infty$ we have the following lower bound valid for any bounded $f \in C^2(\mathbb{R}_+)$ such that $V_A' < \infty$ and $V_L' < \infty$*

$$V_A' - V_L' \geq -\frac{1}{2} V_A \|f''\|_\infty (C_A + C_L).$$

Proof. This immediately follows from Equation (4.7). □

We can improve the inequality in the previous proposition under additional assumptions. If

$$\int_{[0,\infty)} (s-x)^+ d(\mu_A - \mu_L)(s) \leq 0 \quad \text{or} \quad \int_{[0,\infty)} (s-x)^+ d(\mu_A - \mu_L)(s) \geq 0 \qquad (4.8)$$

for every $x \geq 0$, then we can use the Mean Value Theorem directly on Equation (4.5) to get

$$V_A' - V_L' = V_A \frac{f''(\zeta_V)}{2}(C_A - C_L).$$

This means that under one of the conditions in Equation (4.8) we are immunized against every shift if we additionally to value and duration matching have a convexity-matched position.

Convex shift functions

We will now look at convex shift functions, and recall that we know from the discussion after Corollary 4.6 that a single liability is immunized against every convex shift. In the general liability case the situation is more complicated. If two random variables X and Y have $E[X] = E[Y]$ the following two claims are equivalent.

(i) $E[f(X)] \leq E[f(Y)]$ for every convex function f such that the expectations are well defined.

(ii) $E|X - a| \leq E|Y - a|$ is true for every $a \in \mathbb{R}$

(see e.g. Hürlimann [3]). It now follows:

Proposition 4.12 *If we have value and duration matching of assets and liabilities then we are immunized against every convex shift with $V_A' < \infty$ and $V_L' < \infty$ if and only if*

$$E|X_A - x| \geq E|X_L - x| \tag{4.9}$$

for every $x \geq 0$.

Equation (4.9) is called the mean absolute deviation (MAD) constraint in the literature. Every convex function $f : [0, \infty) \to \mathbb{R}$ can be written

$$f(s) = f(0) + f_+'(0)s + \int_{(0,\infty)} (s - x)^+ \gamma(dx), \tag{4.10}$$

where f_+' is the right-hand derivative and γ is the second derivative measure of f. The second-derivative measure of a convex function is a positive Radon measure on $[0, \infty)$ and if f'' exists then $\gamma(dx) = f''(x)dx$. If the assets and liabilities have equal value and duration, then

$$V_A' - V_L' = V_A \left(E\left[\int_{(0,\infty)} (X_A - x)^+ \gamma(dx)\right] - E\left[\int_{(0,\infty)} (X_L - x)^+ \gamma(dx)\right]\right) \tag{4.11}$$

for every shift function such that the expectations exist. We can use the two inequalities

$$(y - x)^+ \geq y - x \quad \text{and} \quad -(y - x)^+ \geq -y,$$

which are valid for every $x \in \mathbb{R}_+$ and $y \in \mathbb{R}$ to get the following proposition.

Proposition 4.13 *Let f be a convex shift function whose second-derivative measure γ fulfills $\int_{(0,\infty)} x\gamma(dx) < \infty$. Then, under value and duration matching and $V_A' < \infty$ and $V_L' < \infty$ we have the lower bound*

$$V_A' - V_L' \geq -V_A \int_{(0,\infty)} x\gamma(dx).$$

Proof. First of all we note that the assumption $\int_{(0,\infty)} x\gamma(dx) < \infty$ implies $\int_{(0,\infty)} \gamma(dx) < \infty$. We have

$$E\left[\int_{(0,\infty)} (X_A - x)^+\gamma(dx)\right] \geq E\left[\int_{(0,\infty)} (X_A - x)\gamma(dx)\right] = D_A \int_{(0,\infty)} \gamma(dx) - \int_{(0,\infty)} x\gamma(dx).$$

and

$$-E\left[\int_{(0,\infty)} (X_L - x)^+\gamma(dx)\right] \geq -E\left[\int_{(0,\infty)} X_L\gamma(dx)\right] = -D_L \int_{(0,\infty)} \gamma(dx).$$

Multiplying these two inequalities with V_A and adding them gives the lower bound. □

Example 4.14 *We want for $a > 0$ to use the convex shift functions (all defined for $s \geq 0$)*

(i) $f(s) = |s - a|$,

(ii) $f(s) = (s - a)^+$, *and*

(iii) $f(s) = (a - s)^+$.

The second-derivative measures are given by

(i) $\gamma(x) = 2\delta_a(x)$,

(ii) $\gamma(x) = \delta_a(x)$, *and*

(iii) $\gamma(x) = -\delta_a(x)$,

where δ_a is the Dirac measure at a. Assuming that we have value and duration matching, the lower bounds will be

(i) $V_A' - V_L' \geq -2aV_A$,

(ii) $V_A' - V_L' \geq -aV_A$, *and*

(iii) $V_A' - V_L' \geq aV_A \geq 0$.

Note that, given value and duration matching, we are always immunized against shifts of the type $f(s) = (a - s)^+$. □

One could argue that the integrability condition demanded in the previous proposition is quite strong. There is, however, one case when it is fulfilled for a large class of shift functions, and that is when the asset and liability measures both have finite support. To see this, assume that neither the assets nor the liabilities generate any cash flows after time $T > 0$. Then we are only interested in shift functions on the interval $[0, T]$, and we can replace infinity in Equation (4.10) with T. This T will then appear in Proposition 4.13, and the integrability condition will be $\int_{(0,T]} x\gamma(dx) < \infty$; which is fulfilled since γ is a Radon measure on $[0, \infty)$.

Assuming, additionally to convexity, that the shift function f is differentiable, we can use the inequality

$$f(x) \geq f(y) + (x - y)f'(y), \quad x, y \in \mathbb{R}_+ \tag{4.12}$$

to get lower bounds.

Remark 4.15 It is well known (see e.g. Appendix A1 in Föllmer & Schied [2]) that for a convex function $f : \mathbb{R} \to \mathbb{R}$

$$f(x) \geq f(y) + (x - y)f'_+(y)$$

holds for every $x \leq y$, and that

$$f(x) \geq f(y) + (x - y)f'_-(y)$$

holds for every $x \geq y$. We want to use the inequality in Equation (4.12) for any $x, y \in \mathbb{R}$, and therefore need to impose a differentiability condition on f.

Theorem 4.16 *Under value and duration matching we have for every convex and differentiable shift function f with derivative bounded by k fulfilling $V'_A < \infty$ and $V'_L < \infty$*

$$V'_A - V'_L \geq -V_A k E|X_A - D_A|.$$

Proof. We will use the inequality

$$-f(X_L) \geq -f(D_A) + (D_A - X_L)f'(X_L), \tag{4.13}$$

where again X_L is a random variable with distribution μ_L. We have

$$E[f(X_L)] = \frac{1}{V_L}\int_{[0,\infty)} f(s)v(s)dL_s = \frac{V'_L}{V_L} < \infty,$$

and since $f'(x)$ is bounded by k we have

$$E|(D_A - X_L)f'(X_L)| \leq (D_A + D_L)k < \infty.$$

This means that we can integrate inequality (4.13) to get

$$-E[f(X_L)] \geq -f(D_A) + E[(D_A - X_L)f'(X_L)].$$

Using Jensen's inequality gives

$$E[f(X_A)] \geq f(E[X_A]) = f(D_A).$$

Finally we arrive at

$$\begin{aligned} V_A' - V_L' &= V_A(E[f(X_A)] - E[f(X_L)]) \\ &\geq -V_A E[(X_L - D_A)f'(X_L)] \\ &\geq -V_A k E|X_A - D_A|. \end{aligned}$$

□

Again, just like in the lower bound given by Corollary 4.7, the lower bound depends on the central absolute deviation. To get this result we need to assume that the convex shift function f is differentiable with bounded derivative, while the lower bound for a single liability only demanded that the change in yield curve was Lipschitz continuous and bounded. Assuming that the shift function f is differentiable gives, expressed in the change of yield,

$$f'(s) = e^{-\Delta'(s)s - \Delta(s)}.$$

In order for this to be bounded, a sufficient condition is not only that the change Δ is bounded, but also that the absolute value of the derivative decreases at least as $1/s$.

Finally we remark that it can be shown that

$$\int_{[0,\infty)} f(s)dA(s) \geq \int_{[0,\infty)} f(s)dL(s) \tag{4.14}$$

holds for every convex function f such that the expectation exists if and only if we have

$$\mu_A = \mu_L Q, \tag{4.15}$$

where

$$\mu_L Q(E) := \int_{[0,\infty)} Q(s, E) d\mu_L(s),$$

and where Q is a stochastic kernel on $\mathbb{R}_+$ fulfilling

$$\int_{[0,\infty)} Q(s, dx) = s \text{ for every } s \in \mathbb{R}_+.$$

(Corollary 2.63 in Föllmer & Schied [2]). Note that (4.15) ⇒ (4.14) follows from Jensen's inequality, and that the reverse implication is the hard one to prove.

4.3 Immunization of stochastic cash flows

The idea of moving from a deterministic to a stochastic model, is of course to be able to model the fact that we may not know at the time of valuation when the liabilities occur, and/or how large they are. When we speak of 'stochastic immunization', we mean that the cash flows generated by the liabilities and/or the assets may be random. We will not assume that the discount function is random. This could, at a first glance, look as a limitation of the model. Note, however, that when we discount, we use the discount factors at the valuation time, and they are perfectly know at this time (as forward rates or as a yield curve). Since we look upon immunization theory as a static theory, when the discount factors are changed this is not due to the fact that time passes, but that we look at a new set of discount factors.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a complete filtered probability space with a filtration fulfilling the usual assumptions of being right-continuous and $\mathcal{F}_0$ containing all the P-null sets of $\mathcal{F}$. A random measure on $\mathbb{R}_+$ is a locally finite kernel ξ from $(\Omega, \mathcal{F})$ to $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, i.e. a mapping

$$\xi : \Omega \times \mathbb{R}_+ \to [0, \infty]$$

such that $\xi(\cdot, A)$ is a locally finite measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ for a.e. $\omega \in \Omega$, and $\xi(\omega, \cdot)$ is a $[0, \infty]$-valued random variable for every $A \in \mathcal{B}(\mathbb{R}_+)$.

Definition 4.17 *By an* increasing process *we mean a finite adapted process whose paths are non-negative, non-decreasing and right-continuous functions on* $[0, \infty)$ *fulfilling* $A_{0-} = A_0 = 0$ *a.s.*

We can think of an increasing process as a special case of a random measure on $[0, \infty)$ in the following sense. For a.e. $\omega \in \Omega$ the path $A_.(\omega)$ is the distribution function of a random measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. We denote this measure $dA_t(\omega)$. If A is an increasing process and X is a non-negative adapted process with paths that are cadlag, then

$$(X \cdot A)_t(\omega) := \int_{[0,t]} X_s(\omega) dA_s(\omega)$$

is a well defined integral. Proposition 3.5 of Chapter I in Jacod & Shiryaev [4] shows that the process $X \cdot A$ is a finite-valued increasing process. With it we associate the measure $d(X \cdot A)_t = X_t dA_t$. Finally, we say that an increasing process is of integrable variation if $E[A_\infty] < \infty$.

We have to choose a suitable model for our stochastic cash flows. One approach would be to assume that the cash flow process is a semimartingale. If the semimartingale has a martingale part consisting of a martingale having paths not of finite variation the integral with respect to this martingale has no 'omega-by-omega' meaning. Due to this we prefer to model the cash flows as an increasing process. One advantage with the approach of modelling cash flow processes as increasing processes is that we can interpret them as random measures on $[0, \infty)$. As in the deterministic case we separate the assets and the liabilities.

Our aim now is to parallel the program outlined in the beginning of Section 4.2. We fix a discount function v and let zero be the time of valuation. The *value* of the liabilities L discounted by v at time 0 is defined by

$$V_L = E\left[\int_{[0,\infty)} v(s) dL_s\right]. \tag{4.16}$$

Since v is deterministic it is trivially adapted. We make the assumption that $V_L < \infty$, a reasonable assumption from an economical point of view, and introduce

$$\widetilde{L}_t = \int_{[0,t]} v(s) dL_s.$$

Note that we have $V_L = E\left[\widetilde{L}_\infty\right] < \infty$. Thus, since v is positive, $\widetilde{L}$ is an increasing process of integrable variation. For liabilities L fulfilling $E\left[\int_{[0,\infty)} v(s) dL_s\right] < \infty$ the *liability measure* μ_L is defined by

$$\mu_L(\omega, dt) = \frac{v(t) dL_t(\omega)}{E\left[\int_{[0,\infty)} v(s) dL_s\right]},$$

which is a random measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. In an analogous way we define V_A, $\widetilde{A}$ and, given that $V_A < \infty$, the *asset measure* μ_A:

$$\mu_A(\omega, dt) = \frac{v(t) dA_t(\omega)}{E\left[\int_{[0,\infty)} v(s) dA_s\right]}.$$

We now choose assets represented by the increasing process A such that

$$V_A = V_L.$$

Again, the shift function f is defined as

$$f(s) = \frac{v'(s)}{v(s)},$$

where v' is a new set of discount factors. We say that a shift function f is *admissible* if

$$E\left[\int_{[0,\infty)} f(s) v(s) dL_s\right] < \infty \quad \text{and} \quad E\left[\int_{[0,\infty)} f(s) v(s) dA_s\right] < \infty.$$

If the shift function f is bounded it is obviously admissible:

$$|V_L'| \le E\left[\left|\int_{[0,\infty)} v_s' dL_s\right|\right] = E\left[\left|\int_{[0,\infty)} f(s) d\widetilde{L}_s\right|\right] \le \|f\|_\infty E|\widetilde{L}|_\infty < \infty,$$

and the same holds when L is replaced by A. We define the *duration* (in this case a random variable) of a stream of cash flows L discounted using the discount function v as

$$D_L(\omega) = \frac{\int_{[0,\infty)} sv(s)dL_s(\omega)}{E\left[\int_{[0,\infty)} v(s)dL_s\right]} = \int_{[0,\infty)} s\frac{v(s)}{V_L}dL_s(\omega) = \int_{[0,\infty)} s\mu_L(\omega, ds).$$

It is assumed here that $0 < V_L < \infty$. We call $E[D_L]$ the *mean duration.* If we have constant discount rate, say $\delta > 0$, then

$$E[D_L]V_L = E\left[\int_{[0,\infty)} se^{-\delta s}dL_s\right] = -\frac{\partial}{\partial\delta}E\left[\int_{[0,\infty)} e^{-\delta s}dL_s\right] = -\frac{\partial}{\partial\delta}V_L(\delta)$$

if interchanging integration and differentiating is allowed. Let

$$m(\delta) = \frac{1}{V_L}E\left[\int_{[0,\infty)} se^{-\delta s}dL_s\right]$$

be the mean duration of the liabilities if the discount rate is δ. We can write

$$m(\delta) = -\frac{\partial V_L/\partial\delta}{V_L} = -\frac{\partial}{\partial\delta}\ln V_L \quad \Leftrightarrow \quad V_L(\delta+\varepsilon) = V_L(\delta)e^{-\int_\varepsilon^{\delta+\varepsilon} m(\delta')d\delta'}.$$

This makes the mean duration equivalent to duration in the deterministic case in the sense that

$$V_L(\delta+\varepsilon) \approx V_L(\delta)\cdot(1 - m(\delta)\varepsilon).$$

Example 4.18 *(Annuity-type cash flows.) Let the cash flows be given by a constant rate of 1 per time unit from 0 until a (not necessarily finite) random time τ. Further let the discount factor have the form $v(s) = e^{-\delta s}$ for some $\delta > 0$. The value V_L of these cash flows are given by*

$$V_L = E\left[\int_0^\tau e^{-\delta s}ds\right] = \frac{1}{\delta}\left(1 - E\left[e^{-\delta\tau}\right]\right) = \frac{1}{\delta}\left(1 - \mathfrak{L}(\delta)\right),$$

where $\mathfrak{L}$ is the Laplace transform of the random time τ. Further

$$\int_0^\tau se^{-\delta s}ds = \frac{1}{\delta^2}\left(\delta\tau e^{-\delta\tau} + 1 - e^{-\delta\tau}\right),$$

so the mean duration is given by

$$E[D_L] = \frac{1}{\delta^2}\cdot\frac{\delta E\left[\tau e^{-\delta\tau}\right] + 1 - E\left[e^{-\delta\tau}\right]}{V_L} = \frac{1}{\delta}\left(1 - \frac{\mathfrak{L}'(\delta)}{\mathfrak{L}(\delta)}\right).$$

□

4.3.1 Single liability immunization

As in the deterministic case we will concentrate on the problem of immunization of a single liability. The time at which the liability is to be paid is a random time (i.e. a random variable $\tau : \Omega \to [0, \infty]$), and the amount to be paid is an a.s. finite and strictly positive random variable. We can write the single liability as

$$L_t(\omega) = K(\omega)\mathbf{1}(t \geq \tau(\omega)),$$

where K is representing the amount to be paid at the random time τ. The value of the single liability is given by

$$V_L = E\left[\int_{[0,\infty)} v(s)dL_s\right] = E\left[Kv(\tau)\right],$$

and its duration by

$$D_L = \frac{\tau K v(\tau)}{E\left[Kv(\tau)\right]}.$$

We will use a general increasing process to model the cash flows generated by the assets used to match the single liability

Proposition 4.19 *Let τ be a random time and K an a.s. finite and strictly positive random variable. If the single liability given by $L_t = K\mathbf{1}(t \geq \tau)$ is financed by assets given by the cash flow A, and we have $V_A = V_L < \infty$ then*

$$V'_A - V'_L \geq -V_A E\left[f(\tau)\frac{Kv(\tau)}{V_L} - f(D_A)\right] \tag{4.17}$$

for every convex shift function f with $V'_A < \infty$ and $V'_L < \infty$. If furthermore $E\left[D_A\right] = E\left[\tau \frac{Kv(\tau)}{V_L}\right]$ then we are immunized against every linear shift.

Proof. We have

$$\begin{aligned} V'_A - V'_L &= E\left[\int_{[0,\infty)} v'(s)dA_s\right] - E\left[Kv'(\tau)\right] \\ &= V_A\left(E\left[\int_{[0,\infty)} f(s)\frac{v(s)}{V_A}dA_s\right] - E\left[f(\tau)\frac{Kv(\tau)}{V_L}\right]\right) \\ &= V_A E\left[\int_{[0,\infty)} f(s)\widetilde{\mu}_A(ds) - f(\tau)\frac{Kv(\tau)}{V_L}\right] \end{aligned}$$

Using Jensen's inequality on the convex function f yields

$$\int_{[0,\infty)} f(s)\widetilde{\mu}_A(ds) \geq f\left(\int_{[0,\infty)} s\widetilde{\mu}_A(ds)\right) = f(D_A).$$

Inserting this in the previous equation yields

$$V_A' - V_L' \geq -V_A E\left[f(\tau)\frac{Kv(\tau)}{V_L} - f(D_A)\right]. \tag{4.18}$$

If f is linear, say $f(s) = a + bs$, then we get

$$\begin{aligned} V_A' - V_L' &\geq V_A E\left[a + bD_A - (a + b\tau)\frac{Kv(\tau)}{V_L}\right] \\ &= V_A\left(a + bE[D_A] - aE\left[\frac{Kv(\tau)}{V_L}\right] - bE\left[\tau\frac{Kv(\tau)}{V_L}\right]\right) \\ &= 0 \end{aligned}$$

since we have $V_A = V_L$ and $E[D_A] = E\left[\tau\frac{Kv(\tau)}{V_L}\right]$. □

If we compare this to the deterministic case, we recall that in that case, we were immunized against every convex shift. We only needed to impose Jensen's inequality to show that a duration matched position made us immune against any convex shift function. In this stochastic setting, however, we can not draw the same conclusion by only invoking Jensen's inequality; a more thorough analysis is needed. Since we know that duration matching is the road to success in the deterministic case, it would be natural to continue in this direction. In the context of stochastic immunization, a duration matched position would be such that

$$D_A = \tau\frac{Kv(\tau)}{V_L} \quad a.s.$$

Thus, we have to impose a.s. equality between two random variables. With a limited (e.g. finite) set of basic assets which we can use to create our asset cash flow stream, it seems too optimistic hoping to find cash flows to fulfill this condition. Instead we can look at positions that are mean duration matched, i.e. positions such that

$$E[D_A] = E\left[\tau\frac{Kv(\tau)}{V_L}\right].$$

Due to the assumptions that $K > 0$ a.s. and $V_L < \infty$, we can interpret $Kv(\tau)/V_L$ as the Radon-Nikodym derivative of a measure $P^\star$ with respect to the initial probability measure P:

$$\frac{dP^\star}{dP} = \frac{Kv(\tau)}{V_L}.$$

We can now write the mean duration matching condition as

$$E[D_A] = E^\star[\tau].$$

This means that mean duration matching of a single liability is equivalent to have the mean asset duration equal to the $P^\star$-*mean* of the time τ. The probability measure $P^\star$ is a purely fictive measure, only making it easy to write the mean duration of a single liability. The problem with mean duration matching is that it only guarantees immunization against linear shifts, and linear shifts have no natural interpretation. We will now shift our attention to positions that are *mean time matched*: $E[D_A] = E[\tau]$. Since mean time matching implies that we are not immunized against linear shifts, we will especially not be immunized against general convex shifts.

Theorem 4.20 *Let τ be a random time and K an a.s. finite and strictly positive random variable. If the single liability given by $L_t = K\mathbf{1}(\tau \le t)$ is financed by assets given by the cash flow A, and we have $V_A = V_L < \infty$, $V'_A < \infty$ and $V'_L < \infty$ then*

$$V'_A - V'_L \ge -V_A E\left[(\tau - E[D_A])f'(\tau)\frac{Kv(\tau)}{V_L}\right] = -V_A E^\star[(\tau - E[D_A])f'(\tau)] \quad (4.19)$$

for every convex and differentiable shift function f. If furthermore $E[D_A] = E[\tau]$ then

$$V'_A - V'_L \ge -V_A Std(\tau)\sqrt{E\left[\left\{f'(\tau)\frac{Kv(\tau)}{V_L}\right\}^2\right]}, \quad (4.20)$$

where $Std(\tau)$ is the standard deviation of τ.

Proof. For f convex and differentiable on $\mathbb{R}_+$ we know that

$$-f(y) \ge -f(x) + (x-y)f'(y)$$

for any $x, y \in \mathbb{R}_+$. The previously cited inequality and Jensen's inequality now yields

$$\begin{aligned} V'_A - V'_L &\ge V_A E\left[f(D_A) - f(\tau)\frac{Kv(\tau)}{V_L}\right] \\ &\ge V_A\left(f(E[D_A]) + E\left[\{-f(E[D_A]) + (E[D_A]-\tau)f'(\tau)\}\frac{Kv(\tau)}{V_L}\right]\right) \\ &= -V_A E\left[(\tau - E[D_A])f'(\tau)\frac{Kv(\tau)}{V_L}\right], \end{aligned}$$

and we have the first claim. We can continue to get

$$\begin{aligned} V'_A - V'_L &= -V_A\left(E\left[(\tau - E[\tau] + E[\tau] - E[D_A])f'(\tau)\frac{Kv(\tau)}{V_L}\right]\right) \\ &\ge -V_A\left(\mathrm{Std}(\tau)\sqrt{E\left[\left\{f'(\tau)\frac{Kv(\tau)}{V_L}\right\}^2\right]} + (E[\tau] - E[D_A])E\left[f'(\tau)\frac{Kv(\tau)}{V_L}\right]\right). \end{aligned}$$

Choosing $E[D_A] = E[\tau]$ gives the claimed lower bound. □

Bounded density

In this section we will assume that the density $dP^\star/dP$ is (uniformly) bounded:

$$\frac{dP^\star}{dP} = \frac{Kv(\tau)}{V_L} \leq B$$

for some constant B. One example when this occurs is when K is deterministic and the yield curve is non-negative. Using the lower bound given by Equation (4.19) we get

$$V_A' - V_L' \geq -V_A B E\left[|\tau - E\left[D_A\right]| \, |f'(\tau)|\right],$$

and using Equation (4.20) we get

$$V_A' - V_L' \geq -V_A B \,\mathrm{Std}(\tau)\sqrt{E\left[f'(\tau)^2\right]}.$$

Example 4.21 *Assume that $K = 1$ and that the discount factor v is bounded from above by 1. Then $dP^\star/dP \leq 1/V_L$, and with $V_A = V_L$ we get the lower bounds*

$$V_A' - V_L' \geq -E\left[|\tau - E\left[D_A\right]| \cdot |f'(\tau)|\right]$$

and

$$V_A' - V_L' \geq -Std(\tau)\sqrt{E\left[f'(\tau)^2\right]}$$

respectively. A constant shift of the yield curve, $f(s) = e^{-\varepsilon s}$, gives

$$V_A' - V_L' \geq -|\varepsilon| E\left[|\tau - E\left[D_A\right]| e^{-\varepsilon\tau}\right]$$

and

$$V_A' - V_L' \geq -|\varepsilon| Std(\tau)\sqrt{E\left[e^{-2\varepsilon\tau}\right]} = -|\varepsilon| Std(\tau)\sqrt{\mathfrak{L}(2\varepsilon)},$$

where $\mathfrak{L}$ is the Laplace transform of the random variable τ. □

Deterministic time

We now assume that the time of the payout is known (i.e. deterministic), but that the amount K to be paid still is unknown. We will denote the time of payout T. The value of this liability is given by

$$V_L = v(T)E\left[K\right].$$

The lower bound for a convex shift f (using Theorem 4.19) becomes

$$\begin{aligned} V_A' - V_L' &\geq V_L E\left[f(D_A) - f(T)\frac{Kv(T)}{V_L}\right] \\ &= v(T)E\left[K\right]\left(E\left[f(D_A)\right] - f(T)\right) \\ &\geq v(T)E\left[K\right]\left\{f(E\left[D_A\right]) - f(T)\right\}. \end{aligned}$$

If the assets are chosen such that $E[D_A] = T$, then we see that we are immunized against every convex shift. This means that with a single liability at a certain time it is enough to be mean duration matched in order to be immunized against every convex shift.

4.3.2 Immunization of general stochastic liabilities

Let us now return to the case where the liabilities are given by an increasing process L. Again we consider only convex shift functions f. Under value and mean duration matching (both assumed being finite) we can use Jensen's inequality and the inequality

$$-f(y) \geq -f(x) + (x-y)f'(y)$$

to get

$$\begin{aligned} V_A' - V_L' &= V_A \left(E\left[\int_{[0,\infty)} f(s)\mu_A(ds) \right] - E\left[\int_{[0,\infty)} f(s)\mu_L(ds) \right] \right) \\ &\geq V_A \left(f(E[D_A]) + E\left[\int_{[0,\infty)} \{-f(E[D_A])(E[D_A]-s)f'(s)\} \mu_L(ds) \right] \right) \\ &= -V_A E\left[\int_{[0,\infty)} (s - E[D_A]) f'(s) \mu_L(ds) \right] \end{aligned}$$

If $\sup_{s\geq 0} |f'(s)| = \|f'\|_\infty < \infty$, then

$$V_A' - V_L' \geq -V_A \|f'\|_\infty E|D_A - E[D_A]|,$$

and we have proved:

Proposition 4.22 *Under value and mean duration matching, we have for every admissible shift function fulfilling* $\|f'\|_\infty < \infty$

$$V_A' - V_L' \geq -V_A \|f'\|_\infty E|D_L - E[D_L]|.$$

Bibliography

[1] Fong, H. G. & Vasicek, O. A. (1984), 'A Risk Minimizing Strategy for Portfolio Immunization', *The Journal of Finance*, Vol. XXXIX, No. 5, 1541-1546

[2] Föllmer, H. & Schied, A. (2002), 'Stochastic Finance. An Introduction in Discrete Time', *de Gruyter Studies in Mathematics* 27

[3] Hürlimann, W (2002), 'On immunization, stop-loss order and the maximum Shiu measure', *Insurance: Mathematics and Economics* 31, 315-325

[4] Jacod J. & Shiryaev, A. N. (1998), 'Limit Theorems for Stochastic Processes', *Springer-Verlag*

[5] Montrucchio, L. & Peccati, L. (1991), 'A note on Shiu-Fisher-Weil immunization theorem', *Insurance: Mathematics and Economics* 10, 125-131

[6] Panjer, H. H. (Ed.) (2001), 'Financial Economics', *The Actuarial Foundation*

[7] Shiu, E. S. W. (1988), 'Immunization of multiple liabilities', *Insurance: Mathematics and Economics* 7, 219-224

[8] Shiu, E. S. W. (1990), 'On Redington's theory of immunization', *Insurance: Mathematics and Economics* 9, 171-175

[9] Uberti, M. (1997), 'A note on Shiu's immunization results', *Insurance: Mathematics and Economics* 21, 195-200

Chapter 5

A Note on Immunized Portfolios, Arbitrage Opportunities and Yield Curve Modelling

5.1 Introduction

Some authors (e.g. Boyle [1], Milgrom [8], Panjer (Ed.) [8], Pedersen et al [11] and Shiu [11]) argue that an immunized position should be seen as an arbitrage opportunity. One can get around this problem by taking a static rather than a dynamic view on immunization. It is then possible to interpret the arbitrage opportunities the authors recognize, as the fact that seemingly innocent modelling of the yield curve allows for arbitrage opportunities. To illustrate, take a market with a continuously compounded constant rate r, i.e. an amount I invested at time t has grown to $Ie^{r(T-t)}$ at time T. Further assume that we have to pay a positive amount K at time $T > 0$ and that we finance this by a stream of cash flows $(A_{t_i})_{i=1,\ldots n}$, where $0 \leq t_1 \leq \ldots t_n < \infty$ and $A_{t_i} > 0$ is the cash flow at time t_i. 'Financed' in this context simply means that the present value of the liability K and the cash flow stream A coincide at time 0:

$$Ke^{-rT} = \sum_{i=1}^{n} e^{-rt_i} A_{t_i}.$$

(By multiplying this relation with $e^{-r\tau}$ we see that the present value of the liability and the assets will coincide at time 0 if and only if they coincide at *any* time $\tau \in \mathbb{R}$.) We will also assume that the cash flow stream (A_{t_i}) is chosen such that

$$TKe^{-rT} = \sum_{i=1}^{n} t_i e^{-rt_i} A_{t_i},$$

i.e. the duration of the two cash flow streams coincide. Now assume that the rate changes from r to r'. The new present value of the assets is given by

$$\begin{aligned}\sum_{i=1}^{n} e^{-r't_i} A_{t_i} &= Ke^{-rT} \sum_{i=1}^{n} e^{-(r'-r)t_i} \frac{A_{t_i} e^{-rt_i}}{Ke^{-rT}} \\ &> Ke^{-rT} \exp\left(-(r'-r) \sum_{i=1}^{n} t_i \frac{A_{t_i} e^{-rt_i}}{Ke^{-rT}} \right) \\ &= Ke^{-r'T},\end{aligned}$$

where we have used Jensen's inequality on the strictly convex function e^{-x} and the fact that the two cash flow streams have the same duration. We can conclude that if the assets have duration equal to the time T at which the liability is due (T is the duration of the liability in this special case), then the present value of the assets is higher than the present value of the liability, *whatever change of the interest rate*, and this is sometimes argued as being an arbitrage opportunity. What really has happened is that we have changed model, leaving the old one (with rate r) for a new one (with rate r'). Another way to see this, is to note that we have not modelled how the interest rate moves. There is no mechanism given (neither endogenously nor exogenously) that tells us what we should expect the interest rates to be tomorrow (or in a week, or in a year, or...). We should bear in mind that immunization theory should be seen as a static theory, not a dynamical. In Section 5.3 we give an example in the Black & Scholes model of a portfolio which increases in value when the constant volatility changes, thus paralleling the case of an immunized position when the constant interest rate changes.

Having concluded that we should not consider an immunized portfolio as an arbitrage opportunity, we move on to the next question: what is the apparent arbitrage appearing in an immunized position? In fact, what some authors have shown is that there are examples of seemingly innocent modelling (we now assume that we actually *model* the change in interest rates) of the interest rates which lead to arbitrage opportunities. Ingersoll et al [4] have shown that if we want to model the yield curve change by shifting the whole curve by a fixed amount, then, in order to avoid arbitrage opportunities, we must assume that the yield curve is flat (i.e. constant over maturity times). In the next section we study the models which allows for a constant yield curve at every valuation time. We will close this note by giving an example in the Black & Scholes model, well known to be free of arbitrage opportunities, which parallels the above mentioned one concerning immunized positions.

5.2 Flat yield curve models

We begin with some definitions. Let $p(t,T)$ be the price at time $t \leq T$ of a zero coupon bond with face value 1 and maturing at time T. We will refer to such a bond as a T-bond. For convenience we will assume that $0 \leq t \leq T < \infty$. By a *flat yield curve model* we will mean a model where $p(t,T)$ can be written

$$p(t,T) = e^{-R(t)\cdot(T-t)} \quad \text{for every } T \geq t,$$

and for some function $R(t)$. That is, fixing a valuation time t, every existing T-bond at that time will be valued using the same discount rate $R(t)$ not depending on T. We will make a distinction between deterministic and stochastic flat yield curve models. In the latter we must define how $R(t)$ (which in the stochastic models also is allowed to be random) is related to the flow of information. The modelling in the stochastic case is performed on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, and the idea is to consider models where $R(t)$ is $\mathcal{F}_t$-measurable for every $t \geq 0$, i.e. R is adapted to the filtration $(\mathcal{F}_t)$.

5.2.1 Deterministic interest rates

The analysis with deterministic interest rates is straightforward, and we need not distinguish between discrete and continuous time. When the interest rates are deterministic, they are completely known over time. If we assume linear pricing, then we must have

$$p(t,T) = p(t,S)p(S,T) \tag{5.1}$$

for any $0 \leq t \leq S \leq T$. To see this we just have to use the facts that 1 unit of currency at time T is worth $p(t,T)$ at time t, and $p(S,T)$ units of currency at time S is worth $p(t,S) \cdot p(S,T)$ at time t. Since the amount $p(S,T)$ at time S is equivalent to 1 unit of currency at T, relation (5.1) must hold. Now take any two times t and T such that $0 \leq t < T$. Then, under a flat yield curve model

$$\begin{aligned} p(0,t)p(t,T) = p(0,T) \quad &\Leftrightarrow \quad R(0)t + R(t)(T-t) = R(0)T \\ &\Leftrightarrow \quad R(t) = R(0). \end{aligned}$$

This shows that under the assumption of linear pricing the only deterministic flat yield curve model is the one with constant interest rate for all $t \geq 0$. As was remarked above, this simple argument above is independent of whether time is discrete or continuous.

5.2.2 Stochastic interest rates

We will only consider continuous time models. A proof in a discrete time setting of the fact that if the interest rate is stochastic, then in order to have a flat yield curve model the interest rate must be constant, can be found in Pedersen et al [9] p. 236.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space which we assume is complete. We also assume that the filtration fulfills the usual conditions of $\mathcal{F}_0$ containing all P-null sets of $\mathcal{F}$ and that the filtration $(\mathcal{F}_t)$ is right-continuous.

Definition 5.1 *A probability measure Q on $(\Omega, \mathcal{F})$ is called an* equivalent martingale measure *if $Q \sim P$ and any discounted price process is a Q-local martingale.*

In a stochastic flat yield curve model we have by definition

$$p(t,T) = e^{-R(t)(T-t)}$$

for some adapted process R. In this stochastic framework the relation given by Equation (5.1) will not hold. This is due to the fact that $p(S,T)$ is not known at time t. We now define (see Björk [3] p. 67 ff. for these definitions) the forward rates $f(t,T)$, for $0 \leq t \leq T$, by

$$f(t,T) = -\frac{\partial \ln p(t,T)}{\partial T},$$

and the short rate $r(t)$, for $t \geq 0$, by

$$r(t) = f(t,t).$$

We immediately see that we must have

$$f(t,T) = R(t), \quad \text{for every} \quad t \geq 0 \quad \text{and} \quad T \geq t,$$

in any stochastic flat yield curve model, and also that

$$R(t) = r(t), \quad t \geq 0 \tag{5.2}$$

in these models.

Proposition 5.2 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space (complete and with the filtration satisfying the usual conditions). The filtration is assumed to be generated by the one-dimensional Brownian motion W. On this space we consider the process r, given by the diffusion*

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW_t; \quad r(0) = r_0,$$

where W is a standard Brownian motion and μ and σ are assumed to be regular enough for the above SDE to possess a unique strong solution. If the price at time t for a T-bond $(T \geq t)$ can be written

$$p(t,T) = E^Q\left[e^{-\int_t^T r(s)ds} \middle| \mathcal{F}_t \right],$$

where Q is some equivalent martingale measure (i.e. the bond market if free of arbitrage), then the only arbitrage free stochastic flat yield curve model is given by $r(t) = r_0$ a.s. for every $t \geq 0$.

Before proving this proposition, we remark that precise conditions on μ and σ for the SDE given in the Proposition to possess a unique solution can be found in Section 5.2 of Karatzas & Shreve [6].

Proof. We recall that having a flat yield curve model means that

$$p(t,T) = e^{-R(t)(T-t)},$$

where R is an adapted process. Using Equation (5.2), which holds in every stochastic flat yield curve model, we can thus write

$$p(t,T) = e^{-r(t)(T-t)}.$$

Defining

$$F^T(t,x) = e^{-x(T-t)}, \tag{5.3}$$

so that $p(t,T) = F^T(t,r(t))$, it is a well known fact from arbitrage pricing (Proposition 3.2 in Björk [3]) that F^T fulfills the PDE

$$\begin{cases} F_t^T + [\mu(t,x) - \lambda(t)\sigma(t,x)]F_x^T + \frac{1}{2}\sigma^2(t,x)F_{xx}^T - xF^T &= 0 \\ F^T(x,T) &= 1 \end{cases}$$

Here $\lambda(t)$ is 'the market price of risk', and it appears in the PDE due to the incompleteness of the model. Inserting F^T from Equation (5.3) in the PDE yields

$$\begin{aligned} 0 &= xF^T(t,x) - [\mu(t,x) - \lambda(t)\sigma(t,x)](T-t)F^T(t,x) \\ &\quad + \frac{1}{2}\sigma^2(t,x)(T-t)^2F^T(t,s) - xF^T(t,x) \\ &= F^T \cdot (T-t)\left[\frac{1}{2}(T-t)\sigma^2(t,x) - (\mu(t,x) - \lambda(t)\sigma(t,x))\right]. \end{aligned}$$

Since $F^T(t,x) > 0$ and $T - t > 0$ as long as $T > t$, the bracket must be identically zero, or:

$$T\sigma^2(t,x) = 2(\mu(t,x) - \lambda(t)\sigma(t,x)) + t\sigma^2(t,x); \quad \text{for every } t \in [0,T) \text{ and } x \in \mathbb{R}.$$

Now, μ and σ only depends on t and x (the short rate has the same drift and diffusion independent of the maturity time of the bond we value), λ depends only on t (the fact that the market price of risk does not depend on T is a result of the absence of arbitrage on the market; see Björk [3] Proposition 3.1), and thus we must have $\sigma = 0$. It now immediately follows that also $\mu = 0$. □

Although this proof is general in the sense that we consider 'any' short rate process, we still have restricted r to be an Itô diffusion. Since every diffusion is a Markov process, we also have imposed that implicit assumption. Finally, the fact that we assume μ and σ to be smooth enough for the SDE governing r should possess a unique strong solution further restricts the short interest rate models covered by Proposition 5.2. It is, however, possible to derive the same result by only assuming that the short rate is a continuous semimartingale. See Section 5.4 for the proof of the following proposition.

Proposition 5.3 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space (complete. and with the filtration satisfying the usual conditions). On this space we consider the process r, which we assume to be a continuous semimartingale with decomposition*

$$r_t = r_0 + M_t + A_t,$$

where r_0 is $\mathcal{F}_0$-measurable, M is a local martingale with $M_0 = 0$ a.s. and A is a process of finite variation with $A_0 = 0$ a.s.. If the price at time t for a T-bond $(T \geq t)$ can be written

$$p(t,T) = E^Q \left[e^{-\int_t^T r(s)ds} \middle| \mathcal{F}_t \right],$$

where Q is some equivalent martingale measure, then the only arbitrage free stochastic flat yield curve model is given by $r(t) = r_0$ a.s. for every $t \geq 0$.

5.3 A concluding example

We will close by giving an example in a very well known context: the Black & Scholes model. It is well known (see e.g. Björk [2] or Karatzas [5]) that this model is free of arbitrage in the sense that there does not exist any self-financing portfolio such that the value process V the portfolio generates satisfies

$$V_0 = 0, \quad V_T \geq 0 \text{ and } P(V_T > 0) > 0$$

for any fixed time $T > 0$. The following example consists of a portfolio with wealth 0 today, and such that its value increases whatever other constant volatility we use. This is not an arbitrage opportunity, since changing volatility in the Black & Scholes approach means that we change model.

Example 5.4 *Consider the Black & Scholes model of a financial market:*

$$\begin{aligned} dB_t &= rB_t dt;\ B_0 = 1 \\ dS_t &= \mu S_t dt + \sigma S_t dW_t;\ S_0 = s, \end{aligned}$$

where the constants fulfill $r > -1$, $\mu \in \mathbb{R}$ and $\sigma > 0$. We will further assume that $s > 0$. Now consider the price at time t of a European call option maturing at time $T > t$ and with strike price X. Write S for the stock price at time t. The price of the option is given by the Black & Scholes formula (Björk [2] p. 90):

$$c = S\Phi(d) - Xe^{-r(T-t)}\Phi\left(d - \sigma\sqrt{T-t}\right),$$

where

$$d = \frac{\ln\left(\frac{S}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

and Φ is the distribution function for a standard normal random variable. We will consider two options at the same time t and with the same maturity time $T > t$ but with different strike prices X_1 and X_2, and study how the price of these options change as the volatility σ changes. Due to this we will often write $d(\sigma)$, suppressing the other arguments. The Vega, denoted by $\mathcal{V}$, is defined as

$$\mathcal{V} = \frac{\partial c}{\partial \sigma},$$

and it can be shown (Björk [2] p. 113) that

$$\mathcal{V}(\sigma) = S\sqrt{T-t}\,\varphi(d(\sigma)) > 0,$$

where $\varphi(x) = \Phi'(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ is the density function of a $N(0,1)$-distributed random variable. We will also need to consider the derivate of the Vega of a European call option with respect to σ (known as the Vomma), which we will denote $\mathcal{W}$. By using the expression of $\mathcal{V}$ given above we get

$$\begin{aligned}
\mathcal{W}(\sigma) &= S\sqrt{T-t}\,\varphi'(d(\sigma))d'(\sigma) \\
&= -S\sqrt{T-t}\,\varphi(d(\sigma))d(\sigma)[d'(\sigma)]^2 \\
&= -\mathcal{V}(\sigma)d(\sigma)[d'(\sigma)]^2.
\end{aligned}$$

Since

$$d'(\sigma) = \frac{\sigma(T-t)\sigma\sqrt{T-t} - \sqrt{T-t}\left[\ln\left(\frac{S}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right]}{\sigma^2(T-t)} = \sqrt{T-t} - \frac{d(\sigma)}{\sigma},$$

we can write

$$\mathcal{W}(\sigma) = -\mathcal{V}(\sigma)d(\sigma)\left[\sqrt{T-t} - \frac{d(\sigma)}{\sigma}\right]^2. \tag{5.4}$$

Now consider a portfolio consisting of 1 European call option with strike price X_1, α European call options with strike price X_2 and finally an amount K in the risk free account. We assume that $X_1 > X_2$. The value at time t of this portfolio is given by (where we emphasize the fact that we consider the value as a function of the volatility)

$$V(\sigma) = c_1(\sigma) + \alpha c_2(\sigma) + K,$$

where c_1 and c_2 is the price at time t of a European call option maturing at time T, and with strike price X_1 and X_2 respectively. With a slight abuse of notation we will now assume that the (constant) volatility is σ_0. To simplify the expressions in the sequel we introduce the short hand notation

$$\mathcal{V}_1^0 = \frac{\partial c_1}{\partial \sigma}(\sigma_0) \quad \textit{and} \quad \mathcal{V}_2^0 = \frac{\partial c_2}{\partial \sigma}(\sigma_0).$$

By choosing

$$\begin{cases} \alpha &= -\mathcal{V}_1^0/\mathcal{V}_2^0 \\ K &= (\mathcal{V}_1^0/\mathcal{V}_2^0)c_2(\sigma_0) - c_1(\sigma_0) \end{cases}$$

we see that

$$V(\sigma_0) = 0 \quad \text{and} \quad V'(\sigma_0) = 0.$$

Thus with the above choice of α and K the value of the portfolio as a function of the constant volatility σ is zero with the present volatility σ_0, and σ_0 is further a stationary point. If $V''(\sigma_0) > 0$ and σ_0 is the only stationary point, then σ_0 is a global minimum point, showing that whatever change of the volatility we consider, the value of the portfolio increases.[1] *We will now show that this is the case. The second derivative of the value of the portfolio with respect to volatility is*

$$V''(\sigma) = \mathcal{V}_1^0 \left(\frac{\mathcal{W}_1(\sigma)}{\mathcal{V}_1^0} - \frac{\mathcal{W}_2(\sigma)}{\mathcal{V}_2^0} \right).$$

Let us introduce the following simplifying notation:

$$\begin{aligned} d_i(\sigma) &= d(\sigma) \quad \text{when the strike price is } X_i \text{ for } i = 1, 2, \text{ and} \\ d_i^0 &= d_i(\sigma_0), \end{aligned}$$

and note that for any volatility σ we have

$$X_1 > X_2 \quad \Rightarrow \quad d_1(\sigma) < d_2(\sigma).$$

Now using expression (5.4) for $\mathcal{W}$ we get

$$\begin{aligned} V''(\sigma_0) &= \mathcal{V}_1^0 \left(d_2^0 \left[\sqrt{T-t} - \frac{d_2^0}{\sigma_0} \right]^2 - d_1^0 \left[\sqrt{T-t} - \frac{d_1^0}{\sigma_0} \right]^2 \right) \\ &= \mathcal{V}_1^0 \frac{d_2^0 - d_1^0}{\sigma_0^2} \left[(d_1^0 + d_2^0 - \sigma_0\sqrt{T-t})^2 + d_1^0 d_2^0 \right] > 0 \end{aligned}$$

(since $X_1 > X_2$ we have $d_1^0 < d_2^0$). To see that σ_0 is the only stationary point, we look at solutions to the equation $V'(\sigma) = 0$. Now,

$$V'(\sigma) = \mathcal{V}_1(\sigma) - \frac{\mathcal{V}_0^1}{\mathcal{V}_0^2}\mathcal{V}_2(\sigma) = S\sqrt{T-t} \left(\varphi(d_1(\sigma)) - \frac{\varphi(d_1^0)}{\varphi(d_2^0)} \varphi(d_2(\sigma)) \right).$$

We remind that $\varphi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$, and thus

$$V'(\sigma) = 0 \quad \Leftrightarrow \quad [d_1(\sigma_0)]^2 - [d_2(\sigma_0)]^2 = [d_1(\sigma)]^2 - [d_2(\sigma)]^2.$$

[1] *We only consider the value $V(\sigma)$ for $\sigma > 0$.*

Now,

$$
\begin{aligned}
[d_1(\sigma)]^2 - [d_2(\sigma)]^2 &= \frac{\left[\ln\left(\frac{S}{X_1}\right) + \left(\frac{\sigma^2}{2} + r\right)(T-t)\right]^2}{\sigma^2(T-t)} \\
&\quad - \frac{\left[\ln\left(\frac{S}{X_2}\right) + \left(\frac{\sigma^2}{2} + r\right)(T-t)\right]^2}{\sigma^2(T-t)} \\
&= \frac{\ln^2\left(\frac{S}{X_1}\right) - \ln^2\left(\frac{S}{X_2}\right) + 2\ln\left(\frac{X_2}{X_1}\right)\left(\frac{\sigma^2}{2} + r\right)(T-t)}{\sigma^2(T-t)},
\end{aligned}
$$

and hence every stationary point to $V(\sigma)$ must fulfill

$$
\frac{\ln^2\left(\frac{S}{X_1}\right) - \ln^2\left(\frac{S}{X_2}\right) + 2\ln\left(\frac{X_2}{X_1}\right)\left(\frac{\sigma^2}{2} + r\right)(T-t)}{\sigma^2(T-t)} = \frac{\ln^2\left(\frac{S}{X_1}\right) - \ln^2\left(\frac{S}{X_2}\right) + 2\ln\left(\frac{X_2}{X_1}\right)\left(\frac{\sigma_0^2}{2} + r\right)(T-t)}{\sigma_0^2(T-t)}.
$$

This equation can be simplified to

$$\sigma^2 = \sigma_0^2,$$

and we see that the only stationary point greater than zero is $\sigma = \sigma_0$. This implies that for $\sigma > 0$, $V(\sigma)$ is minimized at σ_0, which was what we wanted to show.

□

5.4 Proof of Proposition 5.3

Let $p(t,T)$ denote the price of a T-bond at time t (with $0 \le t \le T$). By definition, under an equivalent martingale measure Q the discounted price process $Z_t^T = p(t,T)/B_t$ should be a local martingale, where $B_t = \exp\left(\int_0^t r(s)ds\right)$. Under a flat yield curve model we have $p(t,T) = e^{-r_t(T-t)}$, so in this case

$$Z_t^T = \exp\left(-r_t(T-t) - \int_0^t r_s ds\right).$$

Now introduce the function $F^T(t,x,y) = e^{-x(T-t)-y}$ and the process $Y_t = \int_0^t r_s ds$. We see that for any fixed $T \in \mathbb{R}$ we have $F^T \in C^{2,2,2}(\mathbb{R},\mathbb{R},\mathbb{R};\mathbb{R})$, and we can thus use the Itô formula on $Z_t^T = F^T(t, r_t, Y_t)$. Note that we have fixed T and consider the process Z^T only on $[0,T]$. Theorem 33 in Chapter II of Protter [10] implies that we for $t \in [0, T]$ can write

$$\begin{aligned} Z_t^T &= Z_0^T + \int_0^t \frac{\partial F^T}{\partial t}(s, X_s, Y_s)ds + \int_0^t \frac{\partial F^T}{\partial x}(s-, X_{s-}, Y_{s-})dr_s \\ &+ \int_0^t \frac{\partial F^T}{\partial y}(s, X_s, Y_s)dY_s + \frac{1}{2}\int_0^t \frac{\partial^2 F^T}{\partial x^2}(s, X_s, Y_s)d[r,r]_s^c, \end{aligned}$$

where we have used the fact that Y is a continuous process of finite variation. Now

$$\frac{\partial F^T}{\partial t} = xF^T, \quad \frac{\partial F^T}{\partial x} = -(T-t)F^T, \quad \frac{\partial F^T}{\partial y} = -F^T \text{ and } \frac{\partial^2 F^T}{\partial x^2} = (T-t)^2 F^T,$$

and by using this and the facts that

$$dY_t = r_t dt, \quad dr_t = dM_t + dA_t \text{ and } d[r,r]_t^c = d[M,M]_t^c,$$

we can simplify to get

$$Z_t^T = Z_0^T - \int_0^t (T-(s-))Z_{s-}^T dM_s + \int_0^t Z_s^T (T-s)\left[\frac{1}{2}(T-s)d[M,M]_s^c - dA_s\right]$$

for $t \in [0, T]$. Now, the first two terms on the right-hand side constitute a local martingale (Z_0^T is trivially a local martingale, and it follows from Theorem 17 in Chapter III of Protter [10] that the stochastic integral also is a local martingale). In order for the last term to be a local martingale, it needs to be identically zero. We see that $Z_t^T \cdot (T-t) \geq 0$ for $t \in [0, T]$ (with $Z_t^T \cdot (T-t) > 0$ for $t \in [0, T)$), so we must have

$$\frac{1}{2}(T-t)d[M,M]_t^c - dA_t = 0 \text{ for every } t \in [0, T).$$

Using the facts that $[M, M]_0^c = A_0 = 0$, we can rewrite this as

$$\frac{T}{2}[M,M]_t^c = \frac{1}{2}\int_0^t s\, d[M,M]_s^c + A_t, \quad \text{for } t \in [0, T). \tag{$\star$}$$

Now, relation ($\star$) should hold for any fixed strictly positive time T. Since T is any fixed strictly positive number, and relation ($\star$) should hold for any T (note that the process Z^T, different for every maturity time T, does not enter ($\star$)), we must have $[M, M]_t^c = 0$ for every $t \geq 0$. To see this, take another time $T' > T$. Then for every $t \in [0, T)$ we have

$$\frac{T}{2}[M,M]_t^c = \frac{T'}{2}[M,M]_t^c.$$

Since $T' > T$ we must have $[M, M]_t^c = 0$ for every $t \in [0, T)$. But T is arbitrary, and by letting $T \to \infty$ we get the desired conclusion. Since $M_0 = 0$ and M is a continuous local martingale this in turn implies that $M = 0$ (Theorem 27 in Chapter II of Protter [10]). It finally follows from ($\star$) that we also must have $A = 0$. Thus, the only dynamics of r consistent with no-arbitrage is $r_t = r_0$ a.s. for every $t \geq 0$.

□

Bibliography

[1] Boyle, P. P. (1978), 'Immunization under Stochastic models of the term structure', *Journal of the Institute of Actuaries* 105, 177-187

[2] Björk, T. (1998), 'Arbitrage Theory in Continuous Time', *Oxford University Press*

[3] Björk, T. (1996), in Runggaldier, W. J. (Editor), 'Financial Mathematics', *Lecture Notes in Mathematics, Springer-Verlag*

[4] Ingersoll, J. E. Jr., Skelton J. & Weil, R. L. (1978), 'Duration Forty Years Later', *Journal of Financial and Quantitative Analysis* Volume 13, Issue 4, Proceedings of Thirteenth Annual Conference of the Western Finance Association, 627-650

[5] Karatzas, I. (1997), 'Lectures on the Mathematics of Finance', *American Mathematical Society*

[6] Karatzas, I. & Shreve, S. E. (1997), 'Brownian Motion and Stochastic Calculus', *Springer-Verlag*, 2nd Edition

[7] Milgrom, P. R. (1985), 'Measuring the Interest Rate Risk, *Transactions of the Society of Actuaries* 37, 241-257; Discussion 259-302

[8] Panjer, H. H. (2001), 'Financial Economics: With Applications to Investments, Insurance and Pensions', *The Actuarial Foundation*

[9] Pedersen, H. W., Shiu, E. S. W. & Thorlacius, A. E. (1989), 'Arbitrage-free Pricing of Interest-rate Contingent Claims', *Transactions of the Society of Actuaries* 41, 231-265; Discussion 267-279

[10] Protter, P. (1990), 'Stochastic Integration and Differential Equations', *Springer-Verlag*

[11] Shiu, E. S. W. (1990), 'On Redington's theory of immunization', *Insurance: Mathematics and Economics* 9, 171-175

Chapter 6

Coherent and Convex Measures of Risk in Portfolio Optimization and Insurance

6.1 Introduction

Let $(\Omega, \mathcal{F})$ be a measurable space and let us think of an outcome $\omega \in \Omega$ as a possible scenario (or state of the world). Given is also a measurable mapping $X : \Omega \to \mathbb{R}$, where $X(\omega)$ is to be interpreted as the amount of money to be paid to (or received by) the holder of the contract X. We want to associate X with a number representing how 'risky' X is. Note that no probability measure is given. We denote by $\mathcal{X}$ the set of all possible contracts. One main idea when coherent measures of risk were introduced in Artzner et al [2], [3] was to map the function X to the real numbers. They only considered finite sample spaces, in which case we can identify the set of random variables with $\mathbb{R}^{|\Omega|}$. Since $\mathbb{R}^{|\Omega|}$ is a finite-dimensional space, the subsequent analysis becomes easy. In fact it follows from well known results in finite-dimensional convex analysis together with the defining properties of coherent risk measures that the risk of any random variable can be written $\sup_{Q\in\mathcal{Q}} E^Q[-X]$, for a closed convex set $\mathcal{Q}$ of probability measures. When we move to a general space Ω, as is done in Delbaen [7], some technical problems arise. If we let $\mathcal{X}$ be the set of measurable and bounded functions $X : \Omega \to \mathbb{R}$, then $\mathcal{X}$ is a Banach space under the norm $\|X\| = \sup_{\omega\in\Omega} |X(\omega)|$. Again we have a representation of the risk as the one in the finite-dimensional case, but we are not guaranteed that the set $\mathcal{Q}$ consists of probability measures. Instead $\mathcal{Q} \subseteq \mathcal{M}_{1,f}$, where $\mathcal{M}_{1,f}$ is the set of finitely additive set functions with total mass 1. In order to get a representation with $\mathcal{Q}$ being a set of probability measures we need to make some additional assumptions. First of all introduce a probability measure P on $(\Omega, \mathcal{F})$. We then let $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$, i.e. $\mathcal{X}$ is the set of essentially bounded random variables. If the coherent risk measure satisfy the Fatou property: if (X_n) is a sequence uniformly bounded by 1, then as

$n \to \infty$

$$X_n \xrightarrow{P} X \text{ implies } \rho(X) \leq \liminf_{n\to\infty} \rho(X_n),$$

then the set $\mathcal{Q}$ is an $L^1(P)$-closed and convex set consisting of probability measures all absolutely continuous with respect to P. The measure P serves as a 'reference measure' (see the remark on p. 6 in Delbaen [7]). A problem with the approach of using L^∞ as the underlying space is that it rules out random variables being e.g. normally or exponentially distributed. Since normally distributed random variables are usual in economics and finance, it seem reasonable to consider spaces $\mathcal{X}$ including them. Two natural extensions come to mind. One is to use the set L^0 of all random variables on $(\Omega, \mathcal{F})$, and another one is to consider $\mathcal{X} = L^p$ for $p \in [1, \infty)$. We will consider the later approach, and refer to Delbaen [7] for risk measures defined on L^0.

We will use the risk measures to study problems arising in portfolio optimization and insurance. Many papers on this subject exists: Acerbi & Simonetti [1], De Giorgi [5], Lemus [15], Rockafellar et al [16],[17] is a sample of these. See also the papers by Denault [8], Fischer [10] and Tasche [18] in which risk capital allocation for portfolios is studied. Let $r = (r_1, \ldots, r_n)$ be the vector of asset returns in a financial market. We assume that the covariance matrix $\mathrm{Var}(r)$ is positive definite. If the returns are elliptically distributed and have finite variance, then every linear combination $a^T r$ for $a \in \mathbb{R}^n$ has the property that

$$a^T r \stackrel{d}{=} a^T E[r] + \sqrt{a^T \mathrm{Var}(r) a} Z,$$

where Z is a symmetric random variable with $E[Z] = 0$ and $\mathrm{Var}(Z) = 1$. Every coherent measure of risk is translation invariant (in the sense that $\rho(X + a) = \rho(X) - a$ for every $X \in \mathcal{X}$ and $a \in \mathbb{R}$) and positively homogeneous. Hence, if we measure the risk of the portfolio $w^T r$ by using a distribution invariant (i.e. $X \stackrel{d}{=} Y \Rightarrow \rho(X) = \rho(Y)$) coherent risk measure ρ we get

$$\rho(w^T r) = -w^T E[r] + \sqrt{w^T \mathrm{Var}(r) w} \rho(Z).$$

The problem of minimizing this with respect to the constraints $w^T \mathbf{1} = 1$ and $w^T E[r] = \bar{r}$, where $\bar{r}$ is a target expected return, will have a solution set depending on the sign on $\rho(Z)$.

$$\text{If } \left\{ \begin{array}{l} \rho(Z) < 0 \\ \rho(Z) = 0 \\ \rho(Z) > 0 \end{array} \right\} \text{ then } \left\{ \begin{array}{l} \text{there is no solution to the problem} \\ \text{every vector satsifying the constraints is a solution} \\ \text{there is a unique solution} \end{array} \right\}$$

One can show (see Proposition 6.5 below) that if Z is a symmetric random variable then $\rho(Z) \geq 0$. In view of this we can conclude that when r is elliptically distributed the minimization problem always has a solution, but this solution may not be unique. The Markowitz problem is to minimize the variance given the same constraints as above. For every $\bar{r} \in \mathbb{R}$ it will have a unique solution. If r is elliptically distributed and $\rho(Z) > 0$, then minimizing the risk given by ρ has the same solution as the Markowitz problem. If on the other hand $\rho(Z) = 0$, then we can only say that the solution to the Markowitz problem is included in the set of minimizers to the problem of minimizing $\rho(w^T r)$. In Embrechts et al [9] it is claimed (Theorem 1) that minimizing a coherent and distribution invariant risk measure given the conditions above is equivalent to the Markowitz problem. This is the case only if $\rho(Z) > 0$. A trivial example when this condition is not satisfied is when $\rho(X) = E[-X]$. The insurance application will use the fact that if ρ is a distribution invariant convex measure of risk defined on L^p, then for $X \in L^p$ we have the following pair of inequalities (Proposition 6.17):

$$-E[X] + \rho(0) \leq \rho(X) \leq K\|X - E[X]\|_p - E[X] + \rho(0).$$

Here K is a constant that only depends on the risk measure ρ. If X is interpreted as a claim faced by an insurance company, then $\pi(X) = \rho(-X)$ is a possible way of defining the premium of the claim. Letting $\pi(0) = 0$ we get the upper bound

$$E[X] + K\|X - E[X]\|_p$$

on $\pi(X)$. If the distribution of X is not known, this upper bound is possible to use as a premium principle.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. The set $\mathcal{X}$ of risks fulfill $\mathcal{X} = L^p(\Omega, \mathcal{F}, P) \equiv L^p$ for some $p \in [1, \infty)$. A risk measure ρ is a functional $\rho : \mathcal{X} \to \mathbb{R}$. We say that ρ is *a convex measure of risk* if it is

- Monotone: $X \geq Y \;\Rightarrow\; \rho(X) \leq \rho(Y)$.
- Translation invariant: $\rho(X + m) = \rho(X) - m$ for every $X \in \mathcal{X}$ and $m \in \mathbb{R}$.
- Convex: $\rho(\alpha X + (1 - \alpha)Y) \leq \alpha\rho(X) + (1 - \alpha)\rho(Y)$ for every $X, Y \in \mathcal{X}$ and $\alpha \in [0, 1]$.

A risk measure additionally fulfilling

- Positive homogeneity: $\rho(\lambda X) = \lambda\rho(X)$ for every $X \in \mathcal{X}$ and $\lambda > 0$.

is said to be *coherent*. Finally, we will only consider risk measures fulfilling

- Distribution invariance: $X \stackrel{d}{=} Y \;\Rightarrow\; \rho(X) = \rho(Y)$.

This means that we consider risk measures as functions from the space of probability distributions rather than from the space of random variables.

Example 6.1 *For every $\gamma > 0$ the risk measure*

$$\rho(X) = \gamma \log E\left[e^{-X/\gamma}\right]$$

is convex and distribution invariant, but not coherent. □

The following representation result is an immediate consequence of Proposition 3.8 in Cheridito et al [4] (see also Frittelli & Rosazza Gianin [12]).

Proposition 6.2 *Let $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, be a convex measure of risk. Then ρ can be represented as*

$$\rho(X) = \sup_{Y \in \mathcal{D}} \big(E\left[Y(-X)\right] + \gamma(Y)\big), \tag{6.1}$$

where, with $1/p + 1/q = 1$,

$$\mathcal{D} = \{Y \in L^q \,|\, Y \geq 0 \text{ and } E\left[Y\right] = 1\}$$

and

$$\gamma : \mathcal{D} \to [-\infty, \infty) \text{ fulfills } \sup_{Y \in \mathcal{D}} \gamma(Y) \in \mathbb{R}.$$

We let

$$\mathcal{Q} = \{Q \in \mathcal{M}_1(\Omega, \mathcal{F}) \,|\, dQ = YdP \text{ for some } Y \in \mathcal{D}\}.$$

Then we can write

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \big(E^Q\left[-X\right] + \widehat{\gamma}(Q)\big),$$

where $\widehat{\gamma}(Q) = \gamma(dQ/dP)$. When ρ is coherent it is also positive homogeneous, implying that $\gamma = 0$ on $\mathcal{D}$. The set

$$\mathcal{A}_\rho = \{X \in \mathcal{X} \,|\, \rho(X) \leq 0\}$$

defines the *acceptance set*. We can get ρ out of $\mathcal{A}_\rho$ by observing that

$$\inf_{m \in \mathbb{R}} \{m + X \in \mathcal{A}_\rho\} = \inf_{m \in \mathbb{R}} \{\rho(X) \leq m\} = \rho(X).$$

In Cheridito et al [4] it is shown that the acceptance set of a convex measure of risk is (norm) closed in L^p. Since every convex measure of risk is translation invariant, the fact that the acceptance set is closed is equivalent to $\{X \in \mathcal{X} \,|\, \rho(X) \leq k\}$ being closed for any $k \in \mathbb{R}$. Hence, we can conclude

Proposition 6.3 *Every convex measure of risk $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, is lower semicontinuous.*

We will now present some elementary facts regarding distribution invariant convex risk measures.

Proposition 6.4 *Let $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, be a distribution invariant convex measure of risk. Then*

$$\rho(X) \geq \rho(E\left[X\right]) = -E\left[X\right] + \rho(0)$$

for every $X \in L^p$.

Proof. The proof consists of making a minor modification in the proof of Lemma 4.45 in Föllmer & Schied [11]. Take any $X \in L^p$ and let Y_k, $k = 1, 2, \ldots$, be independent copies of X. Then

$$X_n = \frac{1}{n}\sum_{k=1}^{n} Y_k \;\rightarrow\; E[X] \text{ in } L^p,$$

and lower semicontinuity implies

$$\rho(E[X]) \leq \liminf_{n\to\infty} \rho(X_n).$$

Convexity and distribution invariance gives

$$\rho(X_n) = \rho\left(\frac{1}{n}\sum_{k=1}^{n} Y_k\right) \leq \frac{1}{n}\sum_{k=1}^{n} \rho(Y_k) = \rho(X),$$

and we are done. □

We say that a random variable X is *symmetric* if $X \stackrel{d}{=} -X$.

Proposition 6.5 *If the two real numbers a and b fulfill $|a| \leq b$ and $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, is a distribution invariant convex measure of risk, then*

$$\rho(aX) \leq \rho(bX)$$

holds for every symmetric random variable X.

Proof. We have for $a \in [-1, 1]$

$$\rho(aX) = \rho\left(\frac{a+1}{2}X + \frac{1-a}{2}(-X)\right) \leq \frac{a+1}{2}\rho(X) + \frac{1-a}{2}\rho(-X) = \rho(X),$$

where we have used convexity and distribution invariance. Since $a/b \leq 1$ we get

$$\rho(aX) = \rho\left(\frac{a}{b}(bX)\right) \leq \rho(bX).$$

□

6.2 Portfolio optimization

We consider a variant of the classical Markowitz model when risk is measured by a distribution invariant convex measure of risk. De Giorgi [5] considers the case where the risk is measured by variance, VaR and CVaR, and essentially when the distribution of the asset returns is assumed to be multinormally distributed.

Rockefeller et al [17] extensively study the optimal portfolio problem when the risk is measured using a deviation measure. A coherent measure of risk ρ satisfying

$$E[-X] < \rho(X) < -\inf X$$

is in one-to-one correspondence with a deviation measure D via

$$D(X) = \rho(X - E[X]) \quad \text{and} \quad \rho(X) = E[-X] + D(X).$$

We will study the coherent case quite briefly, focusing on a case where the returns have the form given by Equation (6.4) below.

We assume a financial market consisting of n risky assets. The vector of asset returns $r = (r_1, \ldots, r_n)$ is an n-dimensional vector with mean vector $\boldsymbol{\mu}$ and covariance matrix V. Hence, in this section *we assume that* $r_i \in L^2$ *for every* $i = 1, \ldots, n$. We further assume that $\boldsymbol{\mu}$ is not parallel to the vector $\mathbf{1}$ of ones and that V is positive definite. Now, given a mean return $\bar{r} \in \mathbb{R}$ we are interested in the problem

$$\mathrm{P}(\bar{r}) \quad \left[\begin{array}{ll} \text{minimize} & \rho(w^T r) \\ \text{subject } to & w^T \mathbf{1} = 1 \\ & w^T \boldsymbol{\mu} = \bar{r}. \end{array} \right.$$

Here ρ is a distribution invariant convex measure of risk. At play in this model are two sides: the choice of risk measure and the choice of the distribution of the returns. Let us introduce $\rho_r : \mathbb{R}^n \to \mathbb{R}$ by $\rho_r(w) := \rho(w^T r)$. It is immediate that ρ_r is a convex function. If ρ is coherent then ρ_r is convex, positively homogeneous and sublinear. Any sublinear function $f : \mathbb{R}^n \to \mathbb{R}$ can be written

$$f(x) = \sup_{y \in Y} x^T y \tag{6.2}$$

for some (nonempty) closed convex set $Y \subseteq \mathbb{R}^n$ (see Section V.3 in Chapter V of Hiriart-Urruty & Lemaréchal [13]). Recall that a *subgradient* of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ is a vector $z \in \mathbb{R}^n$ such that

$$f(y) \geq f(x) + z^T (y - x)$$

holds for every $y \in \mathbb{R}^n$. The *subdifferential* of f at x, denoted $\partial f(x)$, is the set of all subgradients at x. It can be shown that for any convex function $f : \mathbb{R}^n \to \mathbb{R}$ it holds that $\partial f(x)$ is a non-empty closed convex set for every x in the domain of f. If f is sublinear, then

$$\partial f(x) = \{ y \in Y \mid f(x) = x^T y \}.$$

Since ρ_r is a (finite) convex function and the constraint set is closed and convex the set of minima is a closed convex subset of the constraint set (Lemma 1.0.1 in Chapter VII of Hiriart-Urruty & Lemaréchal [13]). Since the constraints are linear equality constraints we have 'basic constraint qualification' (BCQ) and we can conclude from Theorem 2.1.4 in Hiriart-Urruty & Lemaréchal [13] that a necessary and sufficient condition for $\widehat{w}$ to be an optimal solution and (λ_1, λ_2) to be a Kuhn-Tucker vector is

a) $1 - \widehat{w}^T \mathbf{1} = 0$ and $\bar{r} - \widehat{w}^T \boldsymbol{\mu} = 0$, and

b) $\mathbf{0} \in (\partial \rho_r(\widehat{w}) - \lambda_1 \mathbf{1} - \lambda_2 \boldsymbol{\mu})$.

Condition b) is equivalent to

$$\lambda_1 \mathbf{1} + \lambda_2 \boldsymbol{\mu} \in \partial \rho_r(\widehat{w}).$$

Not every convex measure of risk will work well in portfolio applications, as is seen in the following example.

Example 6.6 *Let r be normally distributed, and assume that the risk measure ρ is coherent. Then*

$$w^T r \stackrel{d}{=} w^T \boldsymbol{\mu} + \sqrt{w^T V w} Z$$

where $Z \sim N(0,1)$. We get

$$\rho(w^T r) = \rho\big(w^T \boldsymbol{\mu} + \sqrt{w^T V w} Z\big) = \sqrt{w^T V w} \rho(Z) - w^T \boldsymbol{\mu}.$$

We see that if $\rho(Z) > 0$, then $P(\bar{r})$ is essentially the Markowitz problem. But what if $\rho(Z) = 0$? In this case, every *vector w fulfilling $w^T \mathbf{1} = 1$ and $w^T \boldsymbol{\mu} = \bar{r}$ is a solution to the problem. Obviously, such a risk measure is not well suited for portfolio optimization.* □

6.2.1 Coherent measures of risk

In this section we will study problem $P(\bar{r})$ when the risk measure ρ is coherent. Using the representation theorem for coherent measures of risk we can write

$$\rho_r(w) = \rho(w^T r) = \sup_{Q \in \mathcal{Q}} E^Q \left[-w^T r\right] = \sup_{Q \in \mathcal{Q}} w^T E^Q \left[-r\right].$$

Letting

$$Y = \{y \in \mathbb{R}^n \mid y = E^Q \left[-r\right] \text{ for some } Q \in \mathcal{Q}\}$$

we have a representation of ρ on the form given in Equation (6.2). Hence the subdifferential of ρ_r at w is given by

$$\partial \rho_r(w) = \{y \in \mathbb{R}^n \mid y = E^Q \left[-r\right] \text{ for some } Q \in \mathcal{Q} \text{ and } \rho_r(w) = w^T y\}. \tag{6.3}$$

Also note the following representation of ρ_r. Let $Z_Q = dQ/dP \in L^2$. Then

$$\begin{aligned} \rho_r(w) &= \sup_{Q \in \mathcal{Q}} [E\left[-Z_Q\right] E\left[w^T \boldsymbol{\mu}\right] + \operatorname{Cov}(Z_Q, -w^T r)] \\ &= -w^T \boldsymbol{\mu} + \sup_{Q \in \mathcal{Q}} w^T \operatorname{Cov}(Z_Q, -r). \end{aligned}$$

To be able to handle the problem posed in Example 6.6, we need to impose some restrictions on the risk measures. We recall that a coherent measure of risk on L^∞ is said to be *relevant* if

$$\rho(-\mathbf{1}_A) > \rho(0)$$

for every set $A \in \mathcal{F}$ with $P(A) > 0$ (Definition 4.27 in Föllmer & Schied [11]). This is equivalent to the following: for $X \in L^\infty$ it holds that

$$X \leq 0 \text{ and } P(X < 0) > 0 \quad \Rightarrow \quad \rho(X) > 0.$$

Since we work with L^p we need some other, but similar definition of relevance. We propose the following.

Definition 6.7 *A convex measure of risk on L^p, $p \in [1, \infty)$, is said to have Property (P) if $\rho(X) > 0$ for every random variable X with $E[X] = 0$ and $P(X \neq 0) > 0$.*

The idea with a risk measure having property (P) is that the risk associated with a non-degenerate random variable is strictly higher than the risk of the expected value. We know from Lemma 6.4 that a distribution invariant coherent measure of risk fulfills $\rho(X) \geq 0$ for every X with $E[X] = 0$. But in order for ρ to be suitable in portfolio optimization we need this inequality to be strict.

Example 6.8 *Let $\rho : L^p \to \mathbb{R}$ be given by*

$$\begin{aligned}\rho(X) &= -E[X] + b\|(X - E[X])^-\|_p \\ &= -E[X] + bE\left[\{(X - E[X])^-\}^p\right]^{1/p}\end{aligned}$$

for some $b \in (0, 1]$. This is a coherent distribution invariant risk measure (see Fischer [10]). It also satisfies Property (P). To see this take $X \in L^p$ with $E[X] = 0$ and $P(X \neq 0) > 0$. Then

$$\rho(X) = bE\left[\{(X - E[X])^-\}^p\right]^{1/p},$$

and since $P(X \neq 0) > 0$ we have $\rho(X) > 0$. By modifying the proof given in Exercise 7 in Delbaen [6] it is easy to show that we have

$$\mathcal{D} = \{Y \in L^q \,|\, 1 + b(U - E[U]) \text{ for some } U \in L^q \text{ with } U \geq 0 \text{ and } \|U\|_q \leq 1\}.$$

□

When we consider L^2 as the space on which the risk measure is defined, property (P) is equivalent to:

$$E[X] = 0 \text{ and } \mathrm{Var}(X) > 0 \;\Rightarrow\; \rho(X) > 0.$$

An important special case occurs when we can write for every $w \in \mathbb{R}^n$

$$w^T r \stackrel{d}{=} w^T \boldsymbol{\mu} + \sqrt{w^T V w} Z, \tag{6.4}$$

where Z is a random variable having $E[Z] = 0$, $\mathrm{Var}(Z) = 1$ and being independent of w. If Z is a symmetric random variable, then r is elliptically distributed. If every linear combination of r has a representation on the form (6.4), then problem $\mathrm{P}(\bar{r})$ is easy to solve.

Proposition 6.9 *Let ρ be a distribution invariant coherent measure of risk. Assume that r has a distribution given by (6.4) for some Z with $E[Z] = 0$ and $Var(Z) = 1$. Then there exists a solution to $P(\bar{r})$, and the optimal portfolio in the Markowitz problem $M(\bar{r})$ is included in the set of minimizers to $P(\bar{r})$. If ρ has property (P), then $P(\bar{r})$ has a unique solution.*

Proof. The representation (6.4) together with distribution invariance and the fact that ρ is a coherent measure of risk implies

$$\rho(w^T r) = \sqrt{w^T V w}\rho(Z) - w^T \boldsymbol{\mu}.$$

Letting $a = 0$ and $b = 1$ in Lemma 6.5 shows that $\rho(Z) \geq 0$. If $\rho(Z) = 0$, then every vector w fulfilling $w^T \mathbf{1} = 1$ and $w^T \boldsymbol{\mu} = \bar{r}$ is a solution to $\mathrm{P}(\bar{r})$; in particular the set of minimizers will include the solution to $\mathrm{M}(\bar{r})$. If ρ has property (P), then $\rho(Z) > 0$, and $\mathrm{P}(\bar{r})$ is equivalent to $\mathrm{M}(\bar{r})$, a problem that is well known to have a unique solution. □

6.2.2 A Capital Asset Pricing Theory

Consider the model of a financial market as above and add to it a risk-less asset with return r_f. We will only consider distribution invariant coherent risk measures in this section. We let w denote the vector of weights invested in the risky assets, and let w_0 denote the weight in the risk-less asset. Then $w_0 = 1 - w^T \mathbf{1}$ and we get

$$\rho(w^T r + w_0 r_f) = \rho(w^T r) + r_f(w^T \mathbf{1} - 1) = \rho_r(w) + r_f w^T \mathbf{1} - r_f.$$

We want to minimize $\rho(w^T r + w_0 r_f)$ given $w^T \mathbf{1} + w_0 = 1$ and $w^T \boldsymbol{\mu} + w_0 r_f = \bar{r}$ for some given expected return $\bar{r}$. This is equivalent to solving

$$\left[\begin{array}{ll} \text{minimize} & \rho_r(w) + r_f w^T \mathbf{1} \\ \text{subject } to & w^T(\boldsymbol{\mu} - r_f \mathbf{1}) = \bar{r} - r_f \end{array}\right.$$

The first order condition is

$$\lambda(\boldsymbol{\mu} - r_f \mathbf{1}) - r_f \mathbf{1} \in \partial \rho_r(\widehat{w}).$$

Since ρ is coherent we have

$$\lambda(\bar{r} - r_f(1 - w_0)) - r_f(1 - w_0) = \rho_r(\widehat{w}).$$

Now *assume* that there exists an optimal vector of weights w_M such that

(i) $w_M^T \mathbf{1} = 1$, and

(ii) $w_M^T \boldsymbol{\mu} =: \bar{r}_M > r_f$.

If the market is in equilibrium we interpret w_M as the market portfolio, i.e. the wealth weighted portfolio of the market's assets. Inserting such a vector w_M in the first order condition and solving for λ yields

$$\lambda = \frac{\rho_r(w_M) + r_f}{\bar{r}_M - r_f}.$$

Now let z be an element of (the non-empty set) $\partial\rho_r(\widehat{w})$. We can write

$$\boldsymbol{\mu} - r_f \mathbf{1} = \frac{1}{\lambda}(z + r_f \mathbf{1}) = \frac{z + r_f \mathbf{1}}{\rho_r(\widehat{w}) + r_f}(\bar{r}_M - r_f),$$

or

$$\mu_i - r_f = \frac{z_i + r_f}{\rho_r(w_M) + r_f}(\bar{r}_M - r_f), \ i = 1, \ldots, n.$$

If the market portfolio is a solution to the problem with expected return $\bar{r}_M$, then we can introduce the generalized beta-value

$$\beta_{iM}^{\rho,z} = \frac{z_i + r_f}{\rho_r(\widehat{w}_M) + r_f}.$$

Lemus [15] assumes that ρ_r is differentiable and arrives at the same expression for the generalized beta. Now let z_M be any element of $\partial\rho_r(w_M)$ (which also is a non-empty set). We know from Equation (6.3) that

$$z_M = E^{Q_M}[-r] \text{ for some } Q_M \in \mathcal{Q} \text{ and } \rho_r(w_M) = E^{Q_M}\left[-w_M^T r\right].$$

We get

$$\lambda(\mu_i - r_f) - r_f = -E^{Q_M}[r_i] = -E[Y_M r_i] = -\mu_i - \mathrm{Cov}(Y_M, r_i),$$

or

$$(\mu_i - r_f)(1 + \lambda) = -\mathrm{Cov}(Y_M, r_i),$$

where $Y_M = dQ_M/dP$. But

$$1 + \lambda = \frac{\rho_r(w_M) + \bar{r}_M}{\bar{r}_M - r_f} = \frac{E\left[-Y_M w_M^T r\right] + E\left[w_M^T r\right]}{\bar{r}_M - r_f} = -\frac{\mathrm{Cov}(Y_M, w_M^T r)}{\bar{r}_M - r_f},$$

and we get

$$\mu_i - r_f = \frac{\mathrm{Cov}(Y_M, r_i)}{\mathrm{Cov}(Y_M, w_M^T r)} \cdot (\bar{r}_M - r_f). \tag{6.5}$$

If we had minimized the variance instead of $\rho(w)$, then we are back in the classical Markowitz approach, and we would get

$$\mu_i - r_f = \frac{\mathrm{Cov}(r_M, r_i)}{\mathrm{Var}(r_M)} \cdot (\bar{r}_M - r_f).$$

The quotient $\mathrm{Cov}(r_i, r_M)/\mathrm{Var}(r_M)$ is the beta-value of r_i with respect to r_M. In equilibrium r_M is the return of the market portfolio, and then the previous equation is the classical Capital Asset Pricing Theory (CAPM).

6.2.3 Elliptically distributed returns

A random vector $X = (X_1, \ldots, X_n)$ has a spherical distribution if for every orthogonal matrix $U \in \mathbb{R}^{n\times n}$

$$UX \stackrel{d}{=} X.$$

We further say that $Y = (Y_1, \ldots, Y_m)$ has an elliptical distribution if

$$Y \stackrel{d}{=} \boldsymbol{\mu} + AX,$$

where $A \in \mathbb{R}^{m\times n}$ and X has an n-dimensional spherical distribution. Now let r be elliptically distributed. Then

$$w^T r \stackrel{d}{=} w^T \mu + \sqrt{w^T V w} Z$$

for some symmetrically distributed random variable Z independent of w. Again we want to solve

$$\mathrm{P}(\bar{r}) \quad \left[\begin{array}{ll} \text{minimize} & \rho(w^T r) \\ \text{subject to} & w^T \mathbf{1} = 1 \\ & w^T \mu = \bar{r} \end{array} \right.$$

where ρ is a distribution invariant convex measure of risk. We have

$$\rho(w^T r) = \rho(w^T \boldsymbol{\mu} + \sqrt{w^T V w} Z) = \rho(\sqrt{w^T V w} Z) - w^T \boldsymbol{\mu}.$$

Example 6.10 *Let us return to the coherent risk measure in Example 6.8:* $\mathcal{X} = L^p$ *and*

$$\rho(X) = -E[X] + b\|(X - E[X])^-\|_p.$$

We will use the fact that for any random variable Z *we have* $Z^- = (|Z| - Z)/2$, *and get*

$$\rho(Z) = bE\left[(Z^-)^2\right]^{1/2} = \frac{b}{\sqrt{2}}\sqrt{1 - E\left[Z|Z|\right]} = \frac{b}{\sqrt{2}}.$$

Thus, if the random vector r *is elliptically distributed then*

$$\rho_r(w) = \rho(w^T r) = -w^T \boldsymbol{\mu} + \frac{b}{\sqrt{2}}\sqrt{w^T V w}.$$

□

Definition 6.11 *A convex measure of risk* $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, *is said to be* strictly scale increasing *if the mapping*

$$\begin{array}{rccc} g: & [0,\infty) & \to & \mathbb{R} \\ & a & \mapsto & \rho(aX) \end{array}$$

is strictly increasing on $[0, \infty)$ *for every symmetric* $X \in L^p$ *with* $P(X \neq 0) > 0$.

Example 6.12 *Since the function* e^{-x} *is strictly convex the distribution invariant convex measure of risk*

$$\rho(X) = \gamma \ln E\left[e^{-X/\gamma}\right],$$

$\gamma > 0$, *has Property (P). It is also strictly scale increasing. To see this we first note that the function*

$$g(a) = \gamma \ln E\left[e^{-aX/\gamma}\right]$$

is continuously differentiable with derivative

$$g'(a) = -\frac{E\left[Xe^{-aX/\gamma}\right]}{E\left[e^{-aX/\gamma}\right]}.$$

We have $g'(0) = 0$ *and since*

$$\begin{aligned} E\left[Xe^{-aX/\gamma}\right] &= E\left[Xe^{-aX/\gamma}\mathbf{1}(X<0)\right] + E\left[Xe^{-aX/\gamma}\mathbf{1}(X>0)\right] \\ &= -E\left[X\left(e^{aX/\gamma} - e^{-aX/\gamma}\right)\mathbf{1}(X>0)\right] \\ &= -2E\left[X\sinh(aX/\gamma)\mathbf{1}(X>0)\right] \\ &< 0 \end{aligned}$$

for every $a > 0$ *and* $\gamma > 0$ *we have* $g'(a) > 0$, *and hence* g *is strictly increasing on* $[0, \infty)$. □

Recall that for every distribution invariant convex measure of risk ρ and every symmetric random variable X the mapping $a \mapsto \rho(aX)$ is increasing (Proposition 6.5). A *coherent* risk measure having property (P) is also strictly scale increasing.

Proposition 6.13 *Assume that r is elliptically distributed with finite variance. Then there exists a solution to the problem $P(\bar{r})$, and the optimal portfolio in the Markowitz problem $M(\bar{r})$ is included in the set of minimizers to $P(\bar{r})$. If ρ is strictly scale increasing, then $P(\bar{r})$ has a unique solution.*

Proof. Problem $\mathrm{P}(\bar{r})$ is equivalent to

$$\left[\begin{array}{ll} \text{minimize} & \rho(\sqrt{w^T V w} Z) \\ \text{subject to} & w^T \mathbf{1} = 1 \\ & w^T \mu = \bar{r}. \end{array}\right.$$

Since $\rho_r(w)$ and $w^T \boldsymbol{\mu}$ are both convex functions of w, so is $\rho(\sqrt{w^T V w} Z)$. Let $G(w) := \rho(\sqrt{w^T V w} Z)$. Since we have BCQ, $\widehat{w}$ is a minimizer to $\mathrm{P}(\bar{r})$ if and only if

$$\mathbf{0} \in \partial G(\widehat{w}) + \lambda_1 \mathbf{1} + \lambda_2 \boldsymbol{\mu}$$

for some multipliers $\lambda_1, \lambda_2 \in \mathbb{R}$. Since the function $a \mapsto \rho(aZ)$ is increasing (Lemma 6.5) the unique solution w_{mv} of $\mathrm{M}(\bar{r})$ belongs to $\partial G(\widehat{w})$. If ρ is strictly scale increasing then minimizing $\rho(\sqrt{w^T V w} Z)$ is equal to minimizing $w^T V w$, and $\mathrm{P}(\bar{r})$ and $\mathrm{M}(\bar{r})$ are equivalent. □

Corollary 6.14 *If $P(\bar{r})$ has a unique solution, then it coincides with the solution to the Markowitz problem $M(\bar{r})$.*

Even though the optimal solutions may be the same in $\mathrm{P}(\bar{r})$ and the Markowitz problem $\mathrm{M}(\bar{r})$, the values at optimum will generally be different. This means that we may get different portfolio frontiers in the two cases.

6.2.4 Conditional Value-at-risk (CVaR)

An α-quantile of a random variable X is any $q_\alpha \in \mathbb{R}$ such that $P(X \leq q_\alpha) \geq \alpha$ and $P(X < q_\alpha) \leq \alpha$. For a general random variable X with finite mean the Conditional Value-at-Risk at level α is given by

$$\begin{aligned} \mathrm{CVaR}_\alpha(X) &= \frac{1}{\alpha}\big(E\left[-X\mathbf{1}(X \leq q_\alpha)\right] + q_\alpha(F(q_\alpha) - \alpha)\big) \\ &= \frac{1}{\alpha} E\left[(q_\alpha - X)^+\right] - q_\alpha. \end{aligned}$$

where q_α is an α-quantile of X.

Remark 6.15 What we call Conditional Value-at-Risk is called Average Value-at-Risk in Föllmer & Schied [11]. It is also sometimes called Expected Shortfall.

Theorem 6.16 *Let $\alpha \in (0,1)$. $CVaR_\alpha$ is a distribution invariant coherent measure of risk with*

$$\mathcal{D} = \{Y \in L^q \,|\, Y \geq 0,\ E[Y] = 1 \text{ and } \|Y\|_\infty \leq 1/\alpha\}.$$

We will follow the proof of Theorem 4.39 in Föllmer & Schied [11] (where it is proven for $\mathcal{X} = L^\infty$).

Proof. Using the definition of CVaR_α and the fact that $0 \leq Y \leq 1/\alpha$ we get for every $Y \in \mathcal{D}$.

$$\mathrm{CVaR}_\alpha(X) - E[-XY] = \frac{1}{\alpha} E[(X - q_\alpha)(\alpha Y - \mathbf{1}(X \leq q_\alpha))] \geq 0.$$

Hence

$$\mathrm{CVaR}_\alpha(X) \geq \sup_{Y \in \mathcal{D}} E[-XY].$$

Now let

$$k = \begin{cases} 0 & \text{if } P(X = q_\alpha) = 0, \text{ and} \\ (\alpha - P(X < q_\alpha))/P(X = q_\alpha) & \text{otherwise,} \end{cases}$$

and define

$$Y_\alpha = \frac{1}{\alpha}\big(\mathbf{1}(X < q_\alpha) + k\mathbf{1}(X = q_\alpha)\big).$$

Obviously $Y_\alpha \in \mathcal{D}$. Furthermore

$$E[-XY_\alpha] = \frac{1}{\alpha} E[-X\mathbf{1}(X < q_\alpha) - kX\mathbf{1}(X = q)].$$

If $P(X = q_\alpha) = 0$ then

$$E[-XY_\alpha] = \frac{1}{\alpha} E[-X\mathbf{1}(X \leq q_\alpha)]$$

and if $P(X = q_\alpha) > 0$ we have

$$\begin{aligned} E[-XY_\alpha] &= \frac{1}{\alpha} E\big[-X\big(\mathbf{1}(X < q_\alpha) + k\mathbf{1}(X = q_\alpha)\big)\big] \\ &= \frac{1}{\alpha} E[-X\mathbf{1}(X \leq q_\alpha)] + q_\alpha(F(q_\alpha) - \alpha). \end{aligned}$$

□

We will now calculate CVaR for a portfolio when the returns r fulfill

$$w^T r \stackrel{d}{=} w^T \boldsymbol{\mu} + \sqrt{w^T V w} Z$$

for every $w \in \mathbb{R}^n$ and some random variable Z with $E[Z] = 0$, $\mathrm{Var}(Z) = 1$ and independent of w. To simplify, we will assume that Z has a continuous distribution given by the distribution function f. In this case

$$\mathrm{CVaR}_\alpha(X) = E[-X \mid X \leq q_\alpha]$$

for the unique α-quantile q_α. Let $q_\alpha(w)$ and z_α denote the (unique) α-quantiles of $w^T r$ and Z respectively. We get

$$\begin{aligned}
\mathrm{CVaR}_\alpha &= -E\left[w^T r | w^T r \leq q_\alpha(w)\right] \\
&= -w^T \boldsymbol{\mu} - \sqrt{w^T V w} E\left[Z | Z \leq z_\alpha\right] \\
&= -w^T \boldsymbol{\mu} + \sqrt{w^T V w}\left(-\frac{1}{\alpha}\int_{-\infty}^{-z_\alpha} x\, f(x) dx\right).
\end{aligned}$$

Let us write

$$k_{Z,\alpha} = -\frac{1}{\alpha}\int_{-\infty}^{-z_\alpha} x f(x) dx.$$

Since $k_{Z,\alpha} > 0$, the portfolio optimization problem $\mathrm{P}(\bar{r})$ is equivalent to the Markowitz problem $\mathrm{M}(\bar{r})$ when the vector r of assets fulfills (6.4) and the risk measure is CVaR. Let μ be a probability measure on $[0,1]$ (with the Borel σ-algebra): $\mu \in \mathcal{M}_1[0,1]$. We can then define

$$\rho_\mu(X) := \int_{[0,1]} \mathrm{CVaR}_\alpha(X) d\mu(\alpha).$$

The functional ρ_μ is a coherent and distribution invariant risk. For $\mathcal{M} \subseteq \mathcal{M}_1[0,1]$ we let

$$\rho_{\mathcal{M}}(X) := \sup_{\mu \in \mathcal{M}} \rho_\mu(X).$$

Again, $\rho_{\mathcal{M}}$ is a distribution invariant coherent risk measure. If condition (6.4) holds we get

$$\rho_\mu(w^T r) = -w^T \boldsymbol{\mu} + \sqrt{w^T V w}\int_{[0,1]} k_{Z,\alpha} d\mu(\alpha)$$

and

$$\rho_{\mathcal{M}}(w^T r) = -w^T \boldsymbol{\mu} + \sqrt{w^T V w}\sup_{\mu \in \mathcal{M}}\int_{[0,1]} k_{Z,\alpha} d\mu(\alpha)$$

respectively.

6.3 Bounds on the risk measure

Assume that there exists an element $m \in L^2$ such that the price $p(X)$ of an asset with payoff X can be written as $p(X) = E[mX]$. We will assume that the stochastic discount factor $m > 0$ a.s. Letting $R = X/p(X)$ denote the gross return we have the pricing relation $E[mR] = 1$. We can rewrite this as $E[m](E[R] - 1/E[m]) = -\mathrm{Cov}(m, R)$. Cauchy-Schwarz inequality yields the bound

$$\left| \frac{E[R] - 1/E[m]}{\sigma(R)} \right| \leq \frac{\sigma(m)}{E[m]}.$$

In the presence of a risk-free rate R_f it will hold that $E[m] = 1/R_f$. In this case the left-hand side of the previous inequality is the Sharpe ratio. Thus, the stochastic discount factor m puts a bound on the possible values on the Sharpe ratio. We will now show that a similar type of bounds holds in the case of a convex risk measure.

Proposition 6.17 *Let $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, be a convex measure of risk. Then*

$$-E[X] + \rho(0) \leq \rho(X) \leq K(\mathcal{Q})\|X - E[X]\|_p - E[X] + \rho(0),$$

where

$$K(\mathcal{Q}) = \min\left(\sup_{Q \in \mathcal{Q}} \|Y_Q\|_q, \sup_{Q \in \mathcal{Q}} \|Y_Q - 1\|_q \right),$$

$\mathcal{Q}$ is the set of probability measures defining ρ and $Y_Q = dQ/dP$, holds for every $X \in L^p$.

Proof. The first inequality is Proposition 6.4. To show the second one we start with Cauchy-Schwarz:

$$E[-XY_Q] = -E[X] + E[-(X - E[X])Y_Q] \leq -E[X] + \|X - E[X]\|_p \|Y_Q\|_q.$$

Observing that $\rho(0) = \sup_{Q \in \mathcal{Q}} \gamma(Y_Q)$ the inequality

$$\begin{aligned} \rho(X) &\leq \sup_{Q \in \mathcal{Q}} E[-Y_Q X] + \sup_{Q \in \mathcal{Q}} \gamma(Y_Q) \\ &\leq \|X - E[X]\|_p \sup_{Q \in \mathcal{Q}} \|Y_Q\|_q - E[X] + \rho(0). \end{aligned}$$

now follows. But since $E[Y_Q] = 1$ for every Y_Q with $Q \in \mathcal{Q}$ we also have

$$E[-XY_Q] = E[-(X - E[X])(Y_Q - 1)] - E[X] \leq \|X - E[X]\|_p \|Y_Q - 1\|_Q - E[X].$$

It follows that

$$\rho(X) \leq \|X - E[X]\|_p \sup_{Q \in \mathcal{Q}} \|Y - 1\|_q - E[X] + \rho(0),$$

and we are done. □

Corollary 6.18 *Let $\rho : L^p \to \mathbb{R}$, $p \in [1, \infty)$, be a convex measure of risk. If $X_n \to a \in \mathbb{R}$ in L^p then $\lim_{n \to \infty} \rho(X_n) = \rho(a) = -a$. In particular, if (X_i) is a sequence of IID random variables with mean μ, then*

$$\lim_{n \to \infty} \rho\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) = -\mu.$$

Proof. By combining the inequalities in the previous proposition we get

$$E[a - X_n] \leq \rho(X_n) - \rho(a) \leq E[a - X_n] + K(\mathcal{Q})\|X_n - E[X_n]\|_p.$$

Letting $n \to \infty$, the first claim follows. From $\frac{1}{n}\sum_{i=1}^n X_i \to \mu$ in L^p we have the second claim. □

Corollary 6.19 *If $p = 2$, then*

$$K(\mathcal{Q}) = \sup_{Q \in \mathcal{Q}} \sigma(Y_Q).$$

Proof. We only need to observe that $\mathrm{Var}(Y) = \|Y - 1\|_2^2 = \|Y\|_2^2 - 1 \leq \|Y\|_2^2$ holds for any $Y \in L^p$. □

We have the following result regarding CVaR.

Proposition 6.20 *Let the probability space be atomless. If the coherent risk measure is $CVaR_\alpha$, with $\alpha \in (0,1)$, and $\mathcal{X} = L^p$, then*

$$K(\mathcal{Q}) = \min\left\{\left(\frac{1}{\alpha}\right)^{1/p}, \left(1 - \alpha + \alpha\left(\frac{1}{\alpha} - 1\right)^q\right)^{1/q}\right\},$$

where $1/p + 1/q = 1$ as usual.

If furthermore $p \in [1,2]$, then $K(\mathcal{Q}) = \left(1 - \alpha + \alpha\left(\frac{1}{\alpha} - 1\right)^q\right)^{1/q}$ for every $\alpha \in (0,1)$.

Proof. In the case of an atomless space, the following set determines CVaR_α

$$\mathcal{D} = \{Y \in L^q \,|\, Y \geq 0,\ \|Y\|_\infty \leq 1/\alpha \text{ and } E[Y] = 1\}.$$

We need to determine the constants in the expression for $K(\mathcal{Q})$. It holds that

$$\sup_{Y \in \mathcal{D}} \|Y\|_q^q = \left(\frac{1}{\alpha}\right)^{q-1},$$

where the maximum is attained for any random variable $Y^\star$ on the form $Y^\star = \frac{1}{\alpha}\mathbf{1}_A$ with $A \in \mathcal{F}$ a set having $P(A) = \alpha$. The same $Y^\star$ will maximize $\|Y - 1\|_q$. Hence

$$\begin{aligned}
\sup_{Y \in \mathcal{D}} \|Y - 1\|_q^q &= \|Y^\star - 1\|_q^q \\
&= E\left[|-\mathbf{1}_{A^c} + (1/\alpha - 1)\mathbf{1}_A|^q\right] \\
&= E\left[(\mathbf{1}_{A^c} + (1/\alpha - 1)\mathbf{1}_A)^q\right] \\
&= E\left[\mathbf{1}_{A^c} + (1/\alpha - 1)^q\mathbf{1}_A\right] \\
&= 1 - \alpha + \alpha(1/\alpha - 1)^q \\
&= (1 - \alpha)\left[1 + (1/\alpha)^{q-1}\right],
\end{aligned}$$

and

$$\|Y^\star - 1\|_q \leq \|Y^\star\|_q \text{ if and only if } (1-\alpha)\left[1 + (1/\alpha)^{q-1}\right] \leq (1/\alpha)^{q-1}.$$

Simplifying this yields

$$\alpha \geq 1 - \alpha^{2-q}.$$

Since $2 - q \leq 0$ if and only if $p \leq 2$ we can conclude that

$$\text{if } p \in [1,2] \text{ it holds that } \|Y^\star - 1\|_q \leq \|Y^\star\|_q \text{ for every } \alpha \geq 0.$$

□

Example 6.21 *When $\rho = CVaR_\alpha$ and $\mathcal{X} = L^2$ we have*

$$K(\mathcal{Q}) = \sqrt{1 - \alpha + \alpha\left(\frac{1}{\alpha} - 1\right)^2} = \sqrt{1/\alpha - 1},$$

and get the bounds

$$-E[X] + \rho(0) \leq \rho(X) \leq -E[X] + \sigma(X)\sqrt{1/\alpha - 1} + \rho(0).$$

□

Example 6.22 *If we want a coherent risk measure depending on the norm we can not use $-E[X] + k\|X\|_p$ for some $k > 0$, since this is not a coherent risk measure. We can, however, impose norm constraints on the set $\mathcal{D}$. Take*

$$\mathcal{D} = \{Y \in L^q \,|\, Y \geq 0,\ \|Y\|_q \leq \beta \text{ and } E[Y] = 1\}$$

for some $\beta > 1$. Then every $Y^\star$ of the form $Y^\star = \beta^p \mathbf{1}_A$ for any $A \in \mathcal{F}$ with $P(A) = \beta^{-p}$ maximizes $\|Y\|_q$ and $\|Y - 1\|_q$. We get

$$\sup_{Y \in \mathcal{D}} \|Y\|_q = \beta \quad \text{and} \quad \sup_{y \in \mathcal{D}} \|Y - 1\|_q = \left[\left((\beta^p - 1)^q - 1\right)\beta^{-p} + 1\right]^{1/q}$$

for this risk measure. Note especially that $p = 2$ gives the bounds

$$-E[X] + \rho(0) \leq \rho(X) \leq -E[X] + \sigma(X)\sqrt{\beta^2 - 1} + \rho(0),$$

and that with $p = 1$ and $\alpha = 1/\beta$ we have $CVaR_\alpha$. □

6.3.1 Applications to insurance

In this final section we will look at some applications of the bounds in an insurance context. For any $X \in \mathcal{X}$ we define $\pi(X) = \rho(-X)$, and interpret $\pi(X)$ as the premium an insured has to pay in order to hedge the risk X. We will assume that $\rho(0) = \pi(0) = 0$. Since we consider X as a claim we will only consider random variables with support on $[0, \infty)$ in this section. The bounds in Proposition 6.17 then becomes

$$E[X] \leq \pi(X) \leq E[X] + K(\mathcal{Q})\|X - E[X]\|_p.$$

Note that the safety loading $\pi(X) - E[X]$ always will be positive. If $p = 2$, then the bounds becomes

$$E[X] \leq \pi(X) \leq E[X] + \sup_{Q \in \mathcal{Q}} \sigma(Y_Q)\sigma(X).$$

The right-hand side is a 'standard deviation principle'. Note that while $-E[X]$ is a coherent measure of risk, the functional $f(X) = -E[X] + k\sigma(X)$ is not a coherent measure of risk for any $k > 0$ since it violates the axiom of monotonicity (see Artzner et al [3] p. 7). If we let $p = 1$ we get

$$E[X] \leq \pi(X) \leq E[X] + \sup_{Q \in \mathcal{Q}} \|Y_Q\|_\infty E|X - E[X]|.$$

The upper bound given by the right-hand side is an 'absolute deviation principle'. We will end by considering two examples showing explicitly how the bounds can be used in determining the premium.

Example 6.23 *Consider an insurance company facing the claims $X_1, X_2, \ldots, X_n$ with the same mean μ and variance σ^2 (both finite). We also assume that the covariance between every X_k and X_ℓ, $k \neq \ell$ is equal to some constant c. To ensure that the covariance matrix is positive semidefinite, we will assume $0 \leq c \leq \sigma^2$. The average claim is equal to $\frac{1}{n}\sum_{k=1}^n X_k$, and the premium Π to be paid by every insured is defined to be*

$$\Pi = \frac{1}{n}\pi\left(\sum_{k=1}^n X_k\right) = \pi\left(\frac{1}{n}\sum_{k=1}^n X_k\right).$$

The problem with this approach is that in order to calculate the right-hand side of the previous relation we need the distribution of the X_k's. To get around this problem, we can use the upper bound. Let us write $c = \varrho\sigma^2$, where $\varrho \in [0,1]$ is the correlation between any distinct X_k and X_ℓ. We get

$$\mu \leq \Pi \leq \mu + \left[\sigma_Q\sqrt{\varrho + \frac{1-\varrho}{n}}\right]\sigma,$$

where we have written $\sigma_Q = \sup_{Q\in\mathfrak{Q}} \sigma(Y_Q)$. *We denote the upper bound with* $\widetilde{\Pi}$. *Using the premium principle* $\Pi = \widetilde{\Pi}$ *will put us on the safe side, a potential problem being that* $\widetilde{\Pi}$ *is so large that nobody wants to pay it as premium. Two special cases are*

(i) *Uncorrelated claims:* $c = 0$. *This gives*

$$\widetilde{\Pi} = \mu + \frac{\sigma_Q}{\sqrt{n}}\sigma.$$

(ii) *A very large number of claims:* $n \to \infty$. *Now we get*

$$\widetilde{\Pi} = \mu + \left[\sigma_Z \sqrt{\varrho}\right]\sigma.$$

Finally, let us assume that $\pi(0)$ *need* not *be equal to 0. The upper bound now becomes*

$$\widetilde{\Pi} = \mu + \sigma_Q \frac{\sigma}{\sqrt{n}} \sqrt{1 + (n-1)\varrho} + \frac{\pi(0)}{n}.$$

Since

$$\begin{aligned} \sqrt{1+(n-1)\varrho} &= \sqrt{(n-1)\varrho}\sqrt{1+\frac{1}{(n-1)\varrho}} \\ &\sim \sqrt{n\varrho}\left(1+\frac{1}{2\sqrt{n\varrho}}-\frac{1}{8n\varrho}\right), \end{aligned}$$

where $f \sim g$ *means* $f(n)/g(n) \to 1$ *as* $n \to \infty$, *we get the following asymptotic expansion for* $\widetilde{\Pi}$:

$$\widetilde{\Pi} \sim \mu + \sigma_Q\sqrt{c} + \frac{\sigma_Q \sigma}{2\sqrt{n}} + \frac{\pi(0) - \sigma_Q\sigma/(8\sqrt{\varrho})}{n}.$$

□

Example 6.24 *Now assume that the claims are taken from a stable distribution. Then*

$$\sum_{k=1}^{n} X_k \stackrel{d}{=} n^{1/\alpha} X + D_n$$

for some $\alpha \in (0, 2]$ *where the* X_k*'s are independent copies of* X. *We will assume the claims to have finite mean, hence restricting us to* $\alpha \in (1, 2]$. *Then*

$$n\mu = n^{1/\alpha}\mu + D_n \;\Rightarrow\; D_n = \mu(n - n^{1/\alpha}).$$

When ρ is a coherent and distribution invariant risk measure (recall $\pi(X) = \rho(-X)$) the premium Π is given by

$$\Pi = \frac{1}{n}\pi\left(\sum_{k=1}^{n} X_k\right) = \mu + (\pi(X) - \mu)n^{(1-\alpha)/\alpha}.$$

In this case the safety loading is the single loading $\pi(X) - \mu$ times the scaling factor $n^{(1-\alpha)/\alpha}$. □

Bibliography

[1] Acerbi, S. & Simonetti, P. (2002), 'Portfolio Optimization with Spectral Measures of Risk', *Abaxbank*, Milano

[2] Artzner, P., Delbaen, F, Eber, J-M. & Heath, D. (1997), Thinking Coherently', RISK 10, November, 68-71

[3] Artzner, P., Delbaen, F, Eber, J-M. & Heath, D. (1999), 'Coherent Measures of Risk', *Mathematical Finance* 9 No. 3 203-228

[4] Cheridito P., Delbaen, F. & Kupper, M. (2004), 'Coherent and convex monetary risk measures for bounded càdlàg processes', *Stochastic Processes and their applications*, p. 1-22

[5] De Giorgi, E. (2002), 'A Note on Portfolio Selections under Various Risk Measures', *Working Paper No. 122*, Working Paper Series, Institute for Empirical Research in Economics University of Zurich.

[6] Delbaen, F. (2000), 'Coherent Risk Measures', *Lecture Notes*

[7] Delbaen, F. (2000), 'Coherent Risk Measures on General Probability Spaces', *Working Paper*

[8] Denault, M. (2001), 'Coherent Allocation of Risk Capital', *Journal of Risk* Vol. 4 No. 1, 1-34

[9] Embrechts, P., McNeil A. and Straumann, D. (2002), 'Correlation and dependence in risk management: properties and pitfalls', in *Risk Management: Value at Risk and Beyond*, Ed. M. A. H. Dempster, *Cambridge University Press*, 176-223

[10] Fischer, T. (2003), 'Risk capital allocation by coherent risk measures based on one-sided moments', *Insurance: Mathematics and Economics*, 32 (1), 135-146

[11] Föllmer H. & Schied A. (2002), 'Stochastic Finance. An Introduction in Discrete Time', *Walter de Gruyter.*

[12] Frittelli, & Rosazza Gianin, E. (2002), 'Putting order in risk measures', *Journal of Banking & Finance* 26, 1473-1486

[13] Hiriart-Urruty, J.-B. & Lemaréchal, G. (1993), 'Convex Analysis and Minimization Algorithms I', *Springer-Verlag*

[14] Landsman, Z. & Valdez, E.A. (2003), 'Tail conditional expectations for elliptical distributions', *North American Actuarial Journal* 7 (4)

[15] Lemus, G. (1999), 'Portfolio optimization with quantile based risk measures', Ph.D. Thesis *Sloan School of Management, MIT*

[16] Rockafellar, R. T., Uryasev, S. & Zabarankin, M. (2003), 'Deviation Measures in Risk Analysis and Optimization', *Research Report # 2002-7*, Department of Industrial and Systems Engineering, University of Florida

[17] Rockafellar, R. T., Uryasev, S. & Zabarankin, M. (2003), 'Portfolio Analysis with General Deviation Measures', *Research Report # 2003-8*, Department of Industrial and Systems Engineering, University of Florida

[18] Tasche, D. (1999), 'Risk contributions and performance measurement', *Working paper Technische Universität München*