**Extending Dynamic Convex Risk Measures
from Discrete Time to Continuous Time:
Convergence Approach**

M. Stadje

Extending Dynamic Convex Risk Measures From Discrete Time to Continuous Time: a Convergence Approach[*]

Mitja Stadje[†]

Eurandom, Eindhoven Technical University

P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Abstract

We present an approach for the transition from convex risk measures in discrete time to their counterparts in continuous time. The aim of this paper is to show that a large class of convex risk measures in continuous time can be obtained as limits of discrete time-consistent convex risk measures. The discrete-time risk measures are constructed from properly rescaled ('tilted') one-period convex risk measures, using a d-dimensional random walk converging to a Brownian motion. Under suitable conditions (covering many standard one-period risk measures) we obtain convergence of the discrete risk measures to the solution of a BSDE, defining a convex risk measure in continuous time, whose driver can then be viewed as the continuous-time analogue of the discrete 'driver' characterizing the one-period risk. We derive the limiting drivers for the semi-deviation risk measure, Value at Risk, Average Value at Risk, and the Gini risk measure in closed form.

Key words: Dynamic convex risk measures; time-consistency; g-expectation; discretization; convergence; special drivers

JEL Classification: D81, G11, G22, G32

IME: IM 10, IM 30, IE12

AMS Subject Classification: 91B16, 91B70, 91B30, 60Fxx

1 Introduction

In this paper we present an approach for the transition from convex risk measures in discrete time to their counterparts in continuous time. This allows us to obtain interesting continuous-time analogues of the many one-period convex risk measures. Consider a position yielding a payoff depending on some random scenario. The position could be a portfolio containing assets and liabilities, a derivative, or an insurance contract. The goal of a risk measure is usually to 'summarize' the information about the position in a single number which should in some form relate to the potential losses of the position. Coherent risk measures (which were later generalized to convex risk measures) are a particular axiomatic class of risk measures for which this number can be interpreted as a minimal capital reserve. (If a coherent/convex risk measure is multiplied by -1 then its value may be viewed as a price.) Coherent (static) risk measure

---

[†]Corresponding Adress: Eurandom, Eindhoven Technical University, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, Tel.: +31 40 2478109, Fax: +31 40 2478190, Email: Stadje@tue.nl

were introduced in Artzner et al. (1997, 1999); they were inspired by the capital adequacy rules laid out in the Basel Accord. The more general concept of a convex risk measure was developed by Föllmer and Schied (2002, 2004) and Frittelli and Rosazza Gianin (2002). For an analysis of insurance premiums and risk measures the reader can consult Wang et al. (1997), Landsman and Sherris (2001), Delbaen (2002) and Goovaerts et al. (2003).

Dynamic risk measures for financial positions, updating the information at every time instance, have often been considered in a discrete-time setting. Of course, the dynamic theory is based on the concepts of static (one-period) risk measuring. In a dynamical context *time-consistency* is a natural assumption to glue together static risk measures. It means that the same risk is assigned to a financial position regardless of whether it is calculated over two time periods at once or in two steps backwards in time. This is in fact equivalent to the property that if an asset X is preferred to an asset Y under all possible scenarios at some time t then X should also have been preferred before t. For a comparison with weaker notions of time-consistency see for instance Roorda and Schumacher (2007). Time-consistent convex risk measures have been discussed in discrete time by Riedel (2004), Roorda et al. (2005), Detlefsen and Scandolo (2005), Cheridito et al. (2006), Cheridito and Kupper (2006), Föllmer and Penner (2006), Ruszczyński and Shapiro (2006b), Bion-Nadal (2006, 2008), Artzner et al. (2007), Roorda and Schumacher (2007) and Jober and Rogers (2008).

Time-consistent convex risk measures can be also studied in a continuous-time setting, see for instance Peng (1997, 2004), Barrieu and El Karoui (2004, 2009), Frittelli and Gianin (2004), Rosazza Gianin (2006), Delbaen (2006), Klöppel and Schweizer (2007), Jiang (2008), and Bion-Nadal (2008). While being well understood in discrete time, modeling risk measures in continuous time is more challenging. However, in many situations information arrives continuously and it seems natural to assume that the agent is allowed to update his capital reserves at any time. An elegant approach is the use of an operator given by the solution of a backward stochastic differential equation (BSDE), the so-called g-expectation; see Barrieu and Peng (1997, 2004), Barrieu and El Karoui (2004, 2009), Coquet et al. (2002), Fritteli and Rosazzi Gianin (2004), Rosazzi Gianin (2006), and Jiang (2008) and the generalization of Bion-Nadal (2008). However, when risk measures are modeled as solutions of BSDEs, the drivers defining the underlying BSDEs are difficult to interpret.

The aim of this paper is to show that a large class of convex risk measures in continuous time can be obtained as limits of some classes of robust discrete time-consistent convex risk measures. We will also prove that without scaling discrete-time risk measures which are generated by a single one-period coherent risk measure blow up when more and more time instances are taken into account, suggesting that it may not be appropriate to use them *without further scaling*, in situations where new information is coming in frequently. Moreover, the risk measures which do converge without scaling will *always* converge to quadratic BSDEs which in a one-dimensional setting corresponds to the entropic risk measure.

We will construct the converging robust discrete risk measures from properly rescaled ('tilted') one-period convex risk measures, using a d-dimensional random walk converging to a Brownian motion. Under suitable conditions (covering many standard one-period convex risk measures) we obtain convergence of the discrete convex risk measures to the solution of a BSDE whose driver can then be viewed as the continuous-time analog of the discrete 'driver' characterizing the one-period risk. We will derive the limiting driver for the semi-deviation risk measure, Value at Risk, Average Value at Risk, and the Gini risk measure in closed form.

In Sections 2-4 we expound the necessary background material and review briefly the re-

quired recent theory on which our approach is based. In Section 5 we present the underlying random walk setting, prove that, for instance, coherent risk measures blow up when extended *without further scaling* and introduce the scaled and tilted discrete-time convex risk measures. In Section 6 they are characterized as solutions of *discrete-time BSDEs*, and their convergence to continuous-time convex risk measures is derived. The explicit form of the limiting drivers for some examples of one-period convex risk measures is determined in the final Section 7.

2 Setup

We fix a finite time horizon $T > 0$. *Financial positions* are represented by random variables $X \in L^p(\Omega, \mathcal{F}_T, P)$ ($L^p(\mathcal{F}_T)$ for short) with $p \in \{2, \infty\}$ on some common probability space with filtration $(\mathcal{F}_t)_{t \in I}$, where $I \subset [0, T]$ is a set of time instances usually including 0 and T, and $X(\omega)$ is the discounted net worth of the position at maturity time t under the scenario ω. Equalities and inequalities between random variables are understood in the P-almost sure sense. Our goal is to quantify the *risk* of X at time t by an $\mathcal{F}_t$-measurable random variable $\rho_t(X)$ for $t \in I$. $\rho_t(X)$ is often interpreted as a capital reserve requirement at time t for the financial position X conditional on the information given by $\mathcal{F}_t$. We call a collection of mappings $\rho_t : L^p(\mathcal{F}_T) \to L^p(\mathcal{F}_t)$, $t \in I$ a *dynamic convex risk measure* if it has the following properties:

- *Normalization*: $\rho_t(0) = 0$.

- *Monotonicity*: If $X, Y \in L^p(\mathcal{F}_T)$ and $X \leq Y$, then $\rho_t(X) \geq \rho_t(Y)$

- *$\mathcal{F}_t$-Cash Invariance*: $\rho_t(X + m) = \rho_t(X) - m$ for $X \in L^p(\mathcal{F}_T)$ and $m \in L^\infty(\mathcal{F}_t)$.

- *$\mathcal{F}_t$-Convexity*: For $X, Y \in L^p(\mathcal{F}_T)$ $\rho_t(\lambda X + (1 - \lambda)Y) \leq \lambda \rho_t(X) + (1 - \lambda)\rho_t(Y)$ for all $\lambda \in L^\infty(\mathcal{F}_t)$ such that $0 \leq \lambda \leq 1$.

- *$\mathcal{F}_t$-Local Property:* $\rho_t(I_A X_1 + I_{A^c} X_2) = I_A \rho_t(X_1) + I_{A^c} \rho_t(X_2)$ for all $X \in L^p(\mathcal{F}_T)$ and $A \in \mathcal{F}_t$.

- *Time-Consistency:* For $X, Y \in L^p(\mathcal{F}_T)$ $\rho_t(X) \leq \rho_t(Y)$ implies $\rho_s(X) \leq \rho_s(Y)$ for all $s, t \in I$ with $s \leq t$.

Normalization guarantees that the null position does not require any capital reserves. If ρ is not normal but satisfies the other axioms then the agent can consider the operator $\rho(X) - \rho(0)$ without changing his preferences. Monotonicity postulates that if in any scenario X pays not more than Y then X should be considered at least as risky as Y. Cash invariance gives the interpretation of $\rho(X)$ as a capital reserve. Convexity, which under cash invariance is equivalent to quasiconvexity, says that diversification should not be penalized. For a further discussion of these axioms see also Artzner et al. (1997, 1999). Note that many similar axioms for premium principles can be found in the literature, see for instance Deprez and Gerber (1985) or Goovaerts et al. (2003, 2004a). The local property implies that if A is $\mathcal{F}_t$-measurable then the agent should know at time t if A has happened and adjust his risk evaluation accordingly. If ρ is not time-consistent then it would be possible that an agent at time s considers the future payoff X more risky than the future payoff Y although he knows that in the future in every possible scenario X will actually turn out to be less risky than Y. Time-consistency excludes

3

this kind of behavior. For the use of corresponding or similar notions of time-consistency see also Duffie and Epstein (1992), Chen and Epstein (2002) and the references given in the introduction. Epstein and Schneider (2003) and Maccheroni et al. (2006) deal with dynamic preferences.

Note that normalization and cash invariance yield that $\rho_T(X) = -X$. Hence, although the monotonicity property is listed as an axiom (as it is traditionally done) it follows from time-consistency. In a setting with only two time instances both notions are equivalent. Furthermore, if $p = \infty$, the local property is implied by monotonicity and cash invariance since

$$I_A \rho_t(X) \overset{(\geq)}{\leq} I_A \rho_t(X I_A \overset{(+)}{-} \|X\|_\infty I_{A^c}) = I_A \rho_t(X I_A).$$

Hence,

$$\rho_t(I_A X_1 + I_{A^c} X_2) = I_A \rho_t(I_A X_1 + I_{A^c} X_2) + I_{A^c} \rho_t(I_A X_1 + I_{A^c} X_2)$$
$$= I_A \rho_t(X_1) + I_{A^c} \rho_t(X_2).$$

Note that due to the normalization, monotonicity and cash invariance, time-consistency is equivalent to the dynamic programming principle:

For every $X \in L^p(\mathcal{F}_T)$: $\rho_s(X) = \rho_s(-\rho_t(X))$ for all $s \leq t$.

Given a dynamic convex risk measure (ρ_s) we define $\rho_{s,t}$ to be equal to ρ_s restricted to $L^p(\mathcal{F}_t)$, i.e., $\rho_{s,t} = \rho_s | L^p(\mathcal{F}_t)$. Then the dynamic programming principle is equivalent to

$$\rho_{s,t}(X) = \rho_{s,u}(-\rho_u(X)) \text{ for } s, u, t \in I \text{ with } s \leq u \leq t.$$

If $I = [0, T]$ then we call $\rho = (\rho_t)_{t \in I}$ a dynamic convex risk measure in continuous time (CCRM). If $I = \{t_0, t_1, \ldots, t_k\}$, where $0 = t_0 < t_1 < \ldots < t_k = T$, we call ρ a dynamic convex risk measure in discrete time (DCRM). Throughout the rest of this paper, $|\cdot|$ will denote the Euclidean norm.

3 Continuous-time convex risk measures

In the case $I = [0, T]$ it is well-known that solutions of backward stochastic differential equations in continuous time (BSDEs) provide CCRMs; see the references listed in the introduction.

Therefore, we start by introducing the definition of a BSDE. As underlying process we take a d-dimensional Brownian motion $W = (W_t)_{t \in [0,T]}$ on $(\Omega, (\mathcal{F}_t)_{t \in [0,T]}, P)$, where $(\mathcal{F}_t)_{t \in [0,T]}$ denotes the standard filtration. The *driver* of the BSDE is a function

$$g : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$$

which is measurable with respect to $\mathcal{P} \otimes \mathbb{B}(\mathbb{R}) \otimes \mathbb{B}(\mathbb{R}^d)$, where $\mathcal{P}$ denotes the predictable σ-algebra, i.e., the σ-algebra generated by the predictable processes, considered as real-valued mappings on $[0, T] \times \Omega$.

Next, fix an $\mathcal{F}_T$-measurable random variable X as 'terminal condition'. A *solution* of the BSDE defined by g and X is a pair (Y_t, Z_t), $0 \leq t \leq T$, of progressively measurable processes with values in $\mathbb{R} \times \mathbb{R}^d$ satisfying

$$Y_t = X + \int_t^T g(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad t \in [0, T] \tag{3.1}$$

4

and

$$\mathbb{E}[(\int_0^T |Z_s|^2 s)^{1/2}] < \infty, \ \mathbb{E}[\sup_{0 \leq t \leq T} |Y_t|^2] < \infty \tag{3.2}$$

where products of vectors are understood as scalar products. To ensure that a unique solution of the above BSDE exists there are two possible assumptions which are usually made:

1. (H1) The standard case (uniformly Lipschitz driver): $X \in L^2$,
   $\mathbb{E}[\int_0^T |g(t,0,0)|^2 dt] < \infty$, and g uniformly Lipschitz continuous with respect to (y,z), i.e., there exists a constant $K > 0$ such that $dP \times dt$ a.s. for all $(y,y_0,z,z_0) \in \mathbb{R}^{2d+2}$

$$|g(t,y,z) - g(t,y_0,z_0)| \leq K(|y-y_0| + |z-z_0|).$$

   Under these assumptions, Pardoux and Peng proved in 1990 the existence and uniqueness of a solution of the BSDE (3.1)-(3.2).

2. (H2) $X \in L^\infty$, $|g(t,\omega,y,z)| \leq K(1 + |y| + |z|^2)$ $\quad dP \times dt$ a.s. and for every $C > 0$ there exists a $\hat{K}$ such that for all $y \in [-C,C]$

$$|\frac{\partial g(t,\omega,y,z)}{\partial y}| \leq \hat{K}(1 + |z|^2) \text{ and } |\frac{\partial g(t,\omega,y,z)}{\partial z}| \leq \hat{K}(1 + |z|) \quad dP \times dt \text{ a.s.}$$

   Under these assumptions Kobylanksi (2000) proved that the BSDE (3.1)-(3.2) has a unique solution (Y_t, Z_t) such that Y is bounded.

Now let g be a driver which is independent of y, convex in z satisfying $g(t,0) = 0$ and such that (H1) (or (H2)) holds. We can define the *operator* Y^g which assigns to every financial position $X \in L^2(\mathcal{F}_T)$ (or $L^\infty(\mathcal{F}_T)$) the first component of the solution of the corresponding BSDE with terminal condition $-X$, say $Y_t^g = Y_t^g(-X)$. It is known that the mapping $X \mapsto Y_t^g(-X)$, $X \in L^p(\mathcal{F}_T)$, is normal, monotone, cash invariant, convex, time-consistent and satisfies the local property, for $p = 2$ (or $p = \infty$). Consequently, for every driver g satisfying (H1) (or (H2)) the operator defined by

$$\rho_t^g(X) = Y_t^g(-X)$$

defines a CCRM with respect to the filtration $(\mathcal{F}_t)_{t \in [0,T]}$ on L^2 (L^∞), (see also Fritelli and Rosazza Gianin (2004), Peng (1997, 2004), Rosazza Gianin (2006), Jiang (2008) and Barrieu and El Karoui (2009)). There are also certain sufficient conditions under which a CCRM is induced as a solution of a BSDE satisfying either (H1) (Coquet et al. (2002), Peng (2004) and Rosazza Gianin (2006)) or (H2) (Hu et al. (2008)).

   Thus, BSDEs provide an abundance of dynamic convex risk measures in continuous time. However, it is very hard to assess the meaning of the function g. From the properties of BSDEs it follows that if $\bar{g}(t,z)$ satisfies (H1) (or (H2)) then $g \leq \bar{g}$ implies that for every X in $L^2(\mathcal{F}_T)$ ($L^\infty(\mathcal{F}_T)$) and any $t \in [0,T]$ we have $\rho_t^g(X) \leq \rho_t^{\bar{g}}(X)$. Hence, choosing a larger function $\bar{g}$ corresponds to using a more conservative risk measures. It is also well known that if $g(t,z) = \gamma|z|^2$ with $\gamma > 0$ then the CCRM is given by the entropic risk measure $\rho_t^g(X) = e_t^\gamma(X) = \gamma \ln (\mathbb{E}[\exp\{-X/\gamma\}|\mathcal{F}_t])$, see for instance Proposition 6.4, Barrieu and El Karoui (2009). Moreover, it is known, see for instance Rosazza Gianin (2006), that in the case of a financial market with a risky asset given by $dS_t = S_t(\nu_t dt + \sigma_t dW_t)$, with processes ν_t and

σ_t satisfying appropriate assumptions, the price of a replicable contingent claim $-X$ at time t corresponds to $\rho_t^g(X)$ with $g(t, z) = -\frac{\nu_t}{\sigma_t}z$.

However, in many other cases the corresponding CCRM risk measure is difficult to interpret. One of the aims of this paper is to show that many CCRMs can be viewed as limits of extensions of standard one-period convex risk measures. A second goal will be to explicitly extend static risk measures whose behavior is well understood (like the ones listed as examples below) to a continuous-time setting.

An important technical tool is the *dual representation* of Y_t^g (Barrieu and El Karoui (2009)). Let G be the conjugate of g with respect to the variable z, that is,

$$G(t, \omega, u) = \operatorname{ess\,sup}_z(uz - g(t, \omega, z)), \quad u \in \mathbb{R}^d.$$

We will usually omit the variable ω in G. For u we insert special processes μ_t, $t \in [0, T]$. We call $(\mu_t)_{t \in [0,T]}$ a BMO if it is a progressively measurable d-dimensional process satisfying $\sup_t \mathbb{E}[\int_t^T |\mu_s|^2 ds | \mathcal{F}_t] \in L^\infty(\mathcal{F}_T)$. Let $\Gamma^\mu(t) = \exp\{\int_0^t \mu_s dW_s - 1/2 \int_0^t |\mu_s|^d s\}$. If $(\mu_s)_{s \in [0,T]}$ is a BMO, Γ^μ is a uniformly integrable martingale (Barrieu and El Karoui (2009), Theorem 7.2 or Kazamaki (1994), Section 3.3). Thus, we can define a probability measure P^μ by setting $\frac{dP^\mu}{dP} = \Gamma^\mu$. By the Girsanov Theorem, the process $W_t - \int_0^t \mu_s ds$ is a Brownian motion under P^μ. We need the following duality (see Theorem 7.4 in Barrieu and El Karoui (2009)):

Theorem 3.1 *Suppose that X is in $L^2(\mathcal{F}_T)$ (or in $L^\infty(\mathcal{F}_T)$) and (H1) (or (H2)) is satisfied. Then we have*

$$\rho_t^g(X) = \operatorname{ess\,sup}_{\mu \in \mathcal{A}} \mathbb{E}^\mu\Big[X - \int_t^T G(s, \mu_s) ds \Big| \mathcal{F}_t\Big], \tag{3.3}$$

where under (H1) $\mathcal{A}$ is the set of progressively measurable d-dimensional processes bounded by K and under (H2) $\mathcal{A}$ is the set of BMOs. Furthermore, in each case there is a $\mu^* \in \mathcal{A}$ for which the essential supremum is attained for every $t \in [0, T]$.

4 Convex risk measures in discrete time

Now we consider the case of discrete time, that is, $I = \{t_0, t_1, \ldots, t_k\}$ where $0 = t_0 < t_1 < \ldots < t_k = T$. We will only consider the space $L^\infty(\mathcal{F}_T)$ in this chapter. However, since later our discrete time filtration will only have finitely many atoms this will not put any restrictions on our results. From the convexity condition, we can derive a dual representation for DCRMs. For this, we first have to introduce for $i = 0, \ldots, k - 1$ the set of one-step transition densities

$$D_{t_{i+1}} = \{\xi \in L_+^1(\mathcal{F}_{t_{i+1}}) \mid \mathbb{E}[\xi | \mathcal{F}_{t_i}] = 1\}.$$

Denote by P^{t_i} the measure P conditioned on $\mathcal{F}_{t_i}$. Every sequence $(\xi_{t_j})_{j=i+1,\ldots,k} \in D_{t_{i+1}} \times D_{t_{i+2}} \times \ldots \times D_{t_k}$ induces a P-martingale

$$M_{t_r}^\xi = \begin{cases} \displaystyle\prod_{j=i+1}^r \xi_{t_j} & \text{if } r \geq i+1 \\ 1 & \text{if } r \leq i \end{cases}$$

6

and a probability measure Q^ξ by $\frac{dQ^\xi}{dP} = M_T^\xi$. We may identify Q^ξ with its density ξ. Set $D = D_{t_1} \times D_{t_2} \times \ldots \times D_{t_k}$. Note that for $Q^\xi \in D$ and an $\mathcal{F}_{t_{i+1}}$-measurable bounded random variable X, $\mathbb{E}[X \prod_{j=i}^{k-1} \xi_{t_{j+1}} | \mathcal{F}_{t_i}]$ is defined P-a.s. whereas $\mathbb{E}_{Q^\xi}[X | \mathcal{F}_{t_i}]$ is defined only Q^ξ-a.s. However, following Cheridito and Kupper (2006) we will use the notation

$$\mathbb{E}_{Q^\xi}[X | \mathcal{F}_{t_i}] := \mathbb{E}[X \prod_{j=i}^{k-1} \xi_{t_{j+1}} | \mathcal{F}_{t_i}] \text{ for } i = 0, \ldots, k.$$

For $Q \in D_{t_{i+1}}$ we define $\mathbb{E}_Q[X | \mathcal{F}_{t_i}]$ similarly. Now let us assume that we have one-period convex risk measures $F_{t_i} : L^\infty(\mathcal{F}_{t_{i+1}}) \to L^\infty(\mathcal{F}_{t_i})$ for $i = 0, \ldots, k-1$. A one period convex risk measure F_{t_i} may be seen as a dynamic risk measure with two time instances t_i and t_{i+1}. However, we have already noted before that in this case time-consistency is redundant. Hence, F_{t_i} is a one period convex risk measure if and only if it satisfies normalization, monotonicity, $\mathcal{F}_{t_i}$-translation invariance, the $\mathcal{F}_{t_i}$-local property, and $\mathcal{F}_{t_i}$-convexity. The interpretation is that the agent is at time t_i and evaluates payoffs with maturity t_{i+1}. Subsequently, we will also call the $(F_{t_i})_{i \in \{0, \ldots, k-1\}}$ the *generators*. Some examples of *generators* are

Examples 4.1   • *Entropic risk measure:*

$$e_{t_i}^\gamma(X) = \gamma \ln\left(\mathbb{E}\left[\exp\{-\frac{X}{\gamma}\} | \mathcal{F}_{t_i} \right] \right), \quad \gamma \in (0, \infty).$$

• *Semi-deviation risk measure:*

$$S_{t_i}^{\lambda,q}(X) = \mathbb{E}[-X | \mathcal{F}_{t_i}] + \lambda ||(X - \mathbb{E}[X | \mathcal{F}_{t_i}])_-||_{t_i,q}, \quad \lambda \in [0,1], \quad q \in [1, \infty)$$

where the L^q-norm is taken with respect to the measure P conditioned on $\mathcal{F}_{t_i}$.

• *Gini risk measure:*

$$V_{t_i}^\theta(X) = \text{ess}\sup_{Q \in D_{t_{i+1}}} \left\{ \mathbb{E}_Q[-X | \mathcal{F}_{t_i}] - \frac{1}{2\theta} C_{t_i}(Q|P) \right\}, \quad \theta > 0$$

where $C_{t_i}(Q|P)$ is the Gini index, that is, for a measure $Q \in D_{t_{i+1}}$ with corresponding conditional density $\xi_{t_{i+1}}$

$$C_{t_i}(Q|P) = \mathbb{E}\left[\left(\frac{dQ}{dP^{t_i}} - 1 \right)^2 | \mathcal{F}_{t_i} \right] = \mathbb{E}\left[(\xi_{t_{i+1}} - 1)^2 | \mathcal{F}_{t_i} \right].$$

• *Value at Risk:*

$$V@R_{t_i}^\alpha(X) = \inf\{m \ \mathcal{F}_{t_i} - measurable \,|\, P[X + m < 0 | \mathcal{F}_{t_i}] \leq \alpha\}, \quad \alpha \in (0,1].$$

• *Average Value at Risk:*

$$AV@R_{t_i}^\alpha(X) = \frac{1}{\alpha} \int_0^\alpha V@R_{t_i}^\lambda(X) d\lambda, \quad \alpha \in (0,1].$$

7

The entropic risk measure is also called the exponential premium in the insurance literature (Goovaerts et al. (2004b)). It is well-known, that its dual is equal to the Kullback-Leibler divergence. The semi-deviation risk measure penalizes negative deviations of X from its mean. The Gini risk measure is closely related to a mean-variance evaluation. The mean-variance operator $\mathbb{E}[-X] + \frac{\theta}{2}\mathrm{Var}(X)$ is not a convex risk measure since it is not monotone. The Gini risk measure is the largest convex risk measure which agrees with the mean-variance risk measure on its domain of monotonicity, see Maccheroni et al. (2004, 2006). From the definition above we can see that measures Q are penalized according to their deviation from the reference measure P^{t_i}.

Of course, contrary to all other examples above, $V@R$ is not a true generator since it is not convex. However, we will also consider it in our analysis since it is the risk measure which is most often used in practice. $V@R$ is one of the cornerstones of the Basel II accord and the Solvency II requirements. It corresponds to the smallest amount of capital a bank or an insurance needs to add to its position and invest in a risk-free asset such that the probability of a negative outcome is kept below α. However, it is clear from its definition that Value at Risk does not capture the size of a loss if it occurs. Average Value at Risk takes the average of all Values at Risk between zero and α. It is, with sometimes slightly different definitions for non-continuous distributions, also often referred to as Conditional Value at Risk, Expected Shortfall or Tail Value at Risk. Denote by $\bar{L}_+(\mathcal{F}_t)$ the set of all $\mathcal{F}_t$-measurable functions from Ω to $[0, \infty]$.

Definition 4.2 *We call a mapping $D_{t_{i+1}} \to \bar{L}_+(\mathcal{F}_{t_i})$ a one-step penalty function if it satisfies*

(i) $\mathrm{ess\,inf}_{\xi \in D_{t_{i+1}}} \phi_{t_i}(\xi) = 0$;

(ii) $\phi_{t_i}(I_A \xi_1 + I_{A^c} \xi_2) = I_A \phi_{t_i}(\xi_1) + I_{A^c}\phi_{t_i}(\xi_2)$ *for all $A \in \mathcal{F}_{t_i}$.*

We define the *penalty function* $\phi^{F_{t_i}}$ of a one-period convex risk measure F_{t_i} on $D_{t_{i+1}}$ as

$$\phi_{t_i}^{F_{t_i}}(\xi) = \mathrm{ess\,sup}_{X \in L^\infty(\mathcal{F}_{t_{i+1}})}\{\mathbb{E}[-X\xi|\mathcal{F}_{t_i}] - F_{t_i}(X)\}.$$

Examples 4.3 *The penalty functions in Examples 4.1 are (see also Föllmer and Schied (2004), and Ruszczyński and Shapiro (2006a))*

- *for the entropic risk measure:*

$$\phi_{t_i}^{e^\gamma}(\xi_{t_{i+1}}) = \gamma\mathbb{E}\Big[\xi_{t_{i+1}} \log\left(\xi_{t_{i+1}}\right)|\mathcal{F}_{t_i}\Big];$$

- *for the semi-deviation risk measure:*

$$\phi_{t_i}^{S_{t_i}^{\lambda,q}}(\xi_{t_{i+1}}) = J_{M_{t_i}^{\lambda,q}}(\xi_{t_{i+1}})$$

  with

$$M_{t_i}^{\lambda,q} = \{\xi \in D_{t_{i+1}} | \xi = 1 + \bar{\xi} - \int \bar{\xi}dP^{t_i} \text{ for some } \bar{\xi} \in L^{q'}(\mathcal{F}_{t_{i+1}}, P^{t_i}),$$
$$\|\bar{\xi}\|_{q'} \le \lambda, \quad \bar{\xi} \ge 0\},$$

  where q' is chosen such that $1/q + 1/q' = 1$ and $J_A(Q)$ is the indicator function which is 0 if $Q \in A$ and infinity otherwise;

- *for the Gini risk measure:*

$$\phi_{t_i}^{V^\theta}(\xi_{t_{i+1}}) = \frac{1}{2\theta}\mathbb{E}\Big[(\xi_{t_{i+1}} - 1)^2|\mathcal{F}_{t_i}\Big].$$

- *for Average Value at Risk:*

$$\phi_{t_i}^{AV@R^\alpha}(\xi_{t_{i+1}}) = J_{M_{t_i}^\alpha}(\xi_{t_{i+1}})$$

where $M_{t_i}^\alpha$ *is the set of all conditional* t_{i+1}*-one-step densities which are bounded by* $1/\alpha$.

Definition 4.4 *For* $0 \le t \le T$*, we call a mapping* $I : L^\infty(\mathcal{F}_T) \to L^\infty(\mathcal{F}_t)$ *continuous from above if* $I(X_n) \to I(X)$ *P-almost surely for every sequence* $(X_n)_{n\ge1}$ *in* $L^\infty(\mathcal{F}_T)$ *that decreases P-almost surely to some* $X \in L^\infty(\mathcal{F}_T)$*. We call a dynamic convex risk measure* $(\rho_t)_{t\in I}$ *on* $L^\infty(\mathcal{F}_T)$ *continuous from above if, for each* $t \in I$*,* ρ_t *is continuous from above.*

Recall that given a DCRM (ρ_{t_j}), $\rho_{t_j,t_{j+1}}$ is defined as ρ_{t_j} restricted to t_{j+1}. For $Q \in D$ with corresponding sequence of one-step densities (ξ) define $\phi_{t_i}(Q) = \phi_{t_i}(\xi_{t_{i+1}})$. Now in the case that $p = \infty$ for a DCRM Theorem 3.4 and Corollary 3.8 of Cheridito and Kupper (2006) give the following representation.

Proposition 4.5 *Let* $I = \{t_0, \ldots, t_k\}$ *where* $0 = t_0 < t_1 < \ldots < t_k = T$*. Suppose that* $\Big(\rho_s : L^\infty(\mathcal{F}_T) \to L^\infty(\mathcal{F}_s)\Big)_{s\in I}$ *is a DCRM which is continuous from above. Then*

$$\rho_{t_i}(X) = \operatorname{ess\,sup}_{Q\in D} \mathbb{E}_Q[-X - \sum_{j=i}^{k-1} \phi_{t_j}(Q)|\mathcal{F}_{t_i}] \tag{4.1}$$

where the ϕ_{t_j} *are the penalty functions with corresponding generators* $\rho_{t_j,t_{j+1}}$*, i.e.,* $\phi_{t_j} = \phi_{t_j}^{\rho_{t_j,t_{j+1}}}$*. On the other hand,* given *penalty functions* ϕ_{t_j}*, with corresponding generators* $(F_{t_j})_{j=0,\ldots,k-1}$*, defining* ρ *by (4.1) always yields a dynamic convex risk measure which is continuous from above, with* $\rho_{t_j,t_{j+1}} = F_{t_j}$ *for* $j = 0, \ldots, k-1$*.*

Hence, starting with convex risk measures like the ones from Example 4.1 we can construct a risk measure by (4.1), which behaves locally like F_{t_j}. Now it is natural to ask what happens if more and more time instances are taken into account. We will present a setting in which DCRMs can be constructed from properly rescaled one-period convex risk measures (with corresponding scaled penalty functions) such that these DCRMs converge to convex risk measures in continuous time. We will obtain classes of CCRMs which may be interpreted as the limits of discrete time-consistent robust extensions of (scaled and tilted) standard one-period convex risk measures.

5 DCRMs on a filtration generated by a Random walk

Let $N \in \mathbb{N}$. Suppose that for every N we are given a finite sequence $0 = t_0^N < t_1^N < t_2^N \ldots < t_{k(N)}^N = T$ satisfying

$$\lim_{N\to\infty} \sup_{j=1,\ldots,k(N)} \Delta t_j^N \to 0$$

9

where $\Delta t_j^N = t_j^N - t_{j-1}^N$. We may assume without loss of generality that for every $N \in \mathbb{N}$, $\sup_{j=1,\ldots,k(N)} \Delta t_j^N \leq 1$. Let $B_j^{N,l}$ be i.i.d. Bernoulli variables so that $P(B_j^{N,l} = \pm 1) = 1/2$, $l = 1,\ldots,d$, $j = 1,\ldots,k(N)$. Consider the d-dimensional random walk $R_t^N = (R_t^{N,1},\ldots,R_t^{N,d})$ which is constant on each of the intervals $[t_i^N, t_{i+1}^N)$ and whose components at time t_i^N are given by

$$R^{N,l}(t_i^N) = \sum_{j=1}^{i} \sqrt{\Delta t_j^N} B_j^{N,l}, \quad i = 0,\ldots,k(N),\ l = 1,\ldots,d.$$

Denote by $\mathcal{F}^N = (\mathcal{F}_t^N)_{t\in[0,T]}$ the filtration generated by the random walk process. As the filtration $\mathcal{F}^N$ is finite the spaces $L^\infty(\mathcal{F}_T^N)$ and $L^2(\mathcal{F}_T^N)$ coincide. As a result we do not need to distinguish anymore the cases $p = 2$ and $p = \infty$ and will subsequently write $L^0(\mathcal{F}_T^N)$. Using Theorem I.2.3 in Kunita and Watanabe (1981) we can, after changing the probability space, assume that there exists a standard Brownian motion W_t such that

$$\sup_{0\leq t\leq T} |R_t^N - W_t| \to 0 \quad \text{in } L^2.$$

For an $\mathcal{F}^N$-adapted process $(U^N)_{t\in\{t_0^N,\ldots,t_k^N\}}$, let $\Delta U_{t_i^N}^N = U_{t_i^N}^N - U_{t_{i-1}^N}^N$ for $i = 1,\ldots,k$. Subsequently, we will omit the index N whenever N is fixed except (to avoid ambiguities) when referring to the filtration.

Our aim is to *extend* certain one-period risk measures to discrete *multi-period* convex risk measures (adapted to the filtration $\mathcal{F}^N$) and then to obtain convergence to *continuous-time* convex risk measures when taking the limit $N \to \infty$. This will give an approximation, and a nice interpretation, for certain CCRMs. Take, for example, the entropic risk measure in continuous time

$$e_t^\gamma(X) = \gamma \ln \left(\mathbb{E}[\exp\{-\frac{X}{\gamma}\}|\mathcal{F}_t]\right), \quad t \in [0,T]$$

and its discrete-time counterpart

$$e_t^{N,\gamma}(X^N) = \gamma \ln \left(\mathbb{E}[\exp\{-\frac{X^N}{\gamma}\}|\mathcal{F}_t^N]\right), \quad t \in [0,T].$$

Of course, e_t^γ and $e_t^{N,\gamma}$ are dynamic convex risk measures. Let X^N be terminal conditions for the Nth DCRM respectively, with $\sup_N ||X^N||_\infty < \infty$. Now we can use the tools of weak convergence of filtrations[1] to conclude that the convergence of X^N to some bounded X in probability implies that

$$\sup_{0\leq t\leq T} |\mathbb{E}[\exp\{-X^N\}|\mathcal{F}_t^N] - \mathbb{E}[\exp\{-X\}|\mathcal{F}_t]| \to 0 \quad \text{in probability as } N \to \infty$$

(see Proposition 2 and the second point of Remark 1 in Coquet et al. (2001)). As the X^N are uniformly bounded, $\exp\{-\frac{X^N}{\gamma}\}$ is uniformly bounded away from zero which implies that

$$\sup_{0\leq t\leq T} |e_t^{N,\gamma}(X^N) - e_t^\gamma(X)| \to 0 \quad \text{in probability as } N \to \infty.$$

---

[1] $\mathcal{F}^N$ converges weakly to $\mathcal{F}$ if for every $A \in \mathcal{F}_T$, $\mathbb{E}[I_A|\mathcal{F}^N]$ converges in the Skohorod J_1-topology in probability to $\mathbb{E}[I_A|\mathcal{F}]$. If the limit is continuous the convergence holds uniformly in t.

So in the case of the entropic risk measure the transition from discrete to continuous time is straightforward. However, most convex risk measures cannot be written as conditional expectations. In particular, for CCRMs that are only implicitly defined as the solutions of BSDEs it is not clear what they really mean. Our goal will be to interpret the notions of risk for these convex risk measures in continuous time by means of the analogous notions of risk for DCRMs whose local behavior is well understood.

From (3.3) we see that the conjugate $G(s, \omega, \mu_s)ds$ plays the role of the penalty functions $(\phi_{t_i}(\xi_{t_{i+1}}))_{i \in \{0,\dots,k-1\}}$ in (4.1) where we identified Q with the corresponding sequence of its conditional one-step densities $(\xi_{t_{i+1}})$. Suppose that (μ_t) is an $\mathbb{R}^d$-valued, deterministic process. Then to determine which discrete time penalty functions $(\phi_{t_i}(\xi_{t_{i+1}}))_{i \in \{0,\dots,k-1\}}$ could relate to a function of the form $G(s, \omega, \mu_s)ds$ in the limit, we have to find out first how $(\mu_s)_{s \in [0,T]}$ relates to $(\xi_{t_{i+1}})_{i \in \{0,\dots,k-1\}}$. Let us consider the random walk with drift $\sum_{t_j \leq t} \mu_{t_j} \Delta t_{j+1}$. Define

$$\xi_{t_{i+1}}^\mu = 1 + \mu_{t_i} \Delta R_{i+1}^N = 1 + \sum_{l=1}^d \mu_{t_i}^l \Delta R_{i+1}^{N,l}, \quad i = 0, \dots, k(N) - 1 \qquad (5.1)$$

and assume that

$$\xi_{t_{i+1}}^\mu > 0, \quad i = 0, \dots, k(N) - 1.$$

Define $\tilde{P}^\mu$ by

$$\frac{d\tilde{P}^\mu}{dP} = \prod_{i=0}^{k(N)-1} \xi_{t_{i+1}}^\mu.$$

It is not difficult to show that $\tilde{P}^\mu$ is the probability measure under which the random walk with drift is a martingale, i.e., for $l = 1, \dots, d$, the one-dimensional processes

$$R_{t_{i+1}}^{N,\mu,l} = R_{t_{i+1}}^{N,l} - \sum_{j:\ t_j \leq t_i} \mu_{t_j}^l \Delta t_{j+1}, \quad i = 0, \dots, k(N) - 1$$

are martingales under $\tilde{P}^\mu$. (This can be checked by proving that for each $l = 1, \dots, d$ the conditional expectation of $\Delta R_{t_{i+1}}^{N,l} - \mu_{t_i}^l \Delta t_{i+1}$, given $\mathcal{F}_{t_i}^N$, is zero under $\xi_{t_{i+1}}^\mu dP$.) Therefore, $\tilde{P}^\mu$ may be interpreted as the discrete-time analogue of P^μ, defined shortly before Theorem 3.1.

Let $B_{i+1}^N = (B_{i+1}^{N,1}, \dots, B_{i+1}^{N,d})$. Now suppose that the one-step penalty functions $(\phi_{t_i}^N)_{i=0,\dots,k(N)-1}$ are homogeneous, that is, $\phi_{t_i}^N(f(B_{i+1}^N))$ is independent of N and i for *any* measurable function $f : \mathbb{R}^d \to \mathbb{R}$. As we have a filtration generated by a binomial random walk with a homogeneous structure, this assumption seems reasonable and in fact is satisfied in all our examples above. Thus, we may omit the N and i and just write $\phi(f(B_1))$ even when passing to the limit. Considering (3.3), and (4.1) and (5.1) we see that if the risk measures in discrete time have the same scaling as in continuous time we must have that $\phi\big(1 + \mu \Delta R_{i+1}^N\big) = \phi\big(1 + \mu B_1 \sqrt{\Delta t_{i+1}}\big)$ is equal to $G(t_i, \omega, \mu)\Delta t_{i+1} + o(\Delta t_{i+1})$. Since Δt_{i+1} gets arbitrarily small this implies $\phi(1) = 0$. Furthermore, the convex function $h_\phi(x) := \phi(1 + xB_1)$, mapping $\mathbb{R}^d$ to $\mathbb{R}$, must satisfy $h_\phi'(0) = 0$. To see that fix $t \in [0, T)$ and let $i(N) = \sup\{j | t_j^N \leq t\}$. Assume that G is continuous in t a.s. and that for every fixed arbitrary $\mu \in \mathbb{R}^d$, we have $\phi\big(1 + \mu B_1 \sqrt{\Delta t_{i+1}^N}\big) = G(t_i^N, \omega, \mu)\Delta t_{i+1}^N + o(\Delta t_{i+1}^N)$ for $N \in \mathbb{N}$ and $i = 0, \dots, k(N) - 1$. Then,

11

as $\lim_N \sup_i \Delta t_i^N = 0$ we have

$$\lim_{N \to \infty} \frac{\phi\big(1 + \mu B_1 \sqrt{\Delta t_{i(N)+1}^N}\big) - \phi(1)}{\Delta t_{i(N)+1}^N} = \lim_{N \to \infty} \frac{\phi_0\big(1 + \mu B_1 \sqrt{\Delta t_{i(N)+1}^N}\big)}{\Delta t_{i(N)+1}^N} = G(t, \omega, \mu).$$

Hence, G is independent of t and $h'_\phi(0) = 0$. Denote the transposed of μ_{t_i} by $\mu_{t_i}^\#$. It follows that

$$\frac{1}{2}\mu_{t_i^N}^\# h''_\phi(0)\mu_{t_i^N}\Delta t_{i+1}^N \approx \phi\big(1 + \mu_{t_i^N} B_{i+1}\sqrt{\Delta t_{i+1}^N}\big) \approx G(\omega, \mu_{t_i^N})\Delta t_i^N.$$

In particular, if $d = 1$ then $G(\mu) = \gamma|\mu|^2$ for a $\gamma \geq 0$, which corresponds to the entropic risk measure. Hence, the arguments above suggest that the only one-period risk measures which may be extended homogeneously to continuous time are those with a corresponding function h_ϕ with $h_\phi(1) = h'_\phi(1) = 0$. Moreover, if the discrete-time risk measures converge, then if $d = 1$ the limit *must be* the entropic risk measure. Furthermore, even in higher dimensions these extensions are not very rich since they are limited to BSDEs driven by purely quadratic drivers. For a study of more general multi-dimensional quadratic BSDEs and their role in indifference evaluation see Frei et al. (2009). The next proposition shows that all one-period coherent risk measures[1] explode in the limit if they are not properly rescaled.

Proposition 5.1 *Suppose that* $(F_{t_i}^N)_{N,i}$ *is a family of homogeneous coherent risk measures, i.e., for every function* $f : \mathbb{R}^d \to \mathbb{R}$, $F_{t_i}^N(f(B_{i+1}^N))$ *is independent of* N *and* i. *Furthermore, suppose that* F^N *is more conservative than a conditional expectation evaluation, i.e., there exists a* $z_0 \in \mathbb{R}^d$ *such that* $F_{t_i}^N(z_0 B_{i+1}^N) > 0$. *For discrete time payoffs* X^N *define*

$$\rho_T^N(X^N) = -X^N \quad and \quad \rho_{t_i}^N(X) = F_{t_i}^N(-\rho_{t_{i+1}}^N(X^N)). \tag{5.2}$$

Then there exists discrete payoffs X^N *converging to a continuous time payoff* X *in* L^2 *such that for all* $t \in [0, T)$ *we have*

$$\rho_t^N(X^N) \overset{N \to \infty}{\to} \infty.$$

The proof will be deferred to an appendix. Therefore, our approach in the sequel will be to additionally *scale* and *tilt* the one-period risk measures and to investigate which continuous time risk measures may be interpreted as their limit. We will see that under scaling and tilting actually all homogenous risk measures can be extended to continuous time. We will first of all scale penalty functions by the factor Δt_i, i.e., we will consider $\phi_{t_i}(\xi^\mu)\Delta t_{i+1}$. Now

$$\phi_{t_i}(\xi^\mu)\Delta t_{i+1} = \phi_{t_i}\big(1 + \mu_{t_i}\sqrt{\Delta t_{i+1}}B_{t_{i+1}}^N\big)\Delta t_{i+1}$$

and $G(t_i, \omega, \mu_{t_i})\Delta t_{i+1}$ both use a different scaling of μ. Thus, they also measure the risk at different scales and it seems unreasonable to assume that they both should in any form correspond in the limit. However, if we consider penalty functions which have the form $\phi_{t_i}(1 + \dfrac{\xi^\mu - 1}{\sqrt{\Delta t_{i+1}}})$, we might obtain a limiting relation of the form

$$\phi_{t_i}(1 + \frac{\xi^\mu - 1}{\sqrt{\Delta t_{i+1}}})\Delta t_{i+1} = \phi_{t_i}\big(1 + \mu_{t_i} B_{t_{i+1}}^N\big)\Delta t_{i+1} \overset{??}{\approx} G(t_i, \omega, \mu_{t_i})\Delta t_{i+1}.$$

---

[1] A one-period convex risk measure F_{t_i} is coherent if $F_{t_i}(\lambda X) = \lambda F_{t_i}(X)$ for all $\lambda \in L_+^\infty(\mathcal{F}_{t_i})$.

Now let us carry out our program: to define a DCRM from one-period convex risk measures in a time-consistent way and to scale the penalty functions by a time-dependent transformation such that in the limit a CCRM is obtained.

For generators $(F_{t_i})_{i=0,\ldots,k-1}$, we introduce $(\phi_{t_i}^{F_{t_i}}(1+\frac{\xi_{t_{i+1}}-1}{\sqrt{\Delta t_{i+1}}})\Delta t_{i+1})_{i\in\{0,\ldots,k-1\}}$ as dynamic penalty functions, where $(\phi_{t_i}^{F_{t_i}})_{i\in\{0,\ldots,k-1\}}$ are the one-period penalty functions of F_{t_i}. Using (4.1) and identifying Q with its conditional densities then leads to the following definition for our DCRM.

Definition 5.2 *For a collection of generators $(F_{t_i})_{i=0,\ldots,k-1}$, with penalty functions $(\phi_{t_i}^{F_{t_i}})_{i=0,\ldots,k-1}$, we define its* robust *extension as*

$$\rho_{t_i}(X) = \sup_{Q\in D} \mathbb{E}_Q\Big[-X - \sum_{j=i}^{k-1} \phi_{t_j}^{F_{t_j}}(1+\frac{\xi_{t_{j+1}}^Q-1}{\sqrt{\Delta t_{j+1}}})\Delta t_{j+1}\Big|\mathcal{F}_{t_i}^N\Big]. \tag{5.3}$$

We call $(\rho_{t_i})_{i=0,\ldots,k}$ defined by (5.3) the robust extended discrete-time convex risk measure.

The above definition scales the penalty function similarly to the continuous time case. The term robustness is motivated in the following way: assume that all one-period risk measures are 'the same', like for example in the case $F_{t_i}^N = AV@R_{t_i}^\alpha$ for all N and all t_i. Then we will see later that if the grid reaches a certain refinement, increasing the number of time instances at which the risk manager recalibrates his risk does not lead to a substantial change of this risk. Without this kind of robustness dynamic risk measurements are not suitable in situations where information is coming in frequently, since slightly different time grids can lead to completely different risk evaluations.

Corollary 5.3 *ρ defined by (5.3) is a DCRM which is continuous from above.*

Proof. Since the filtration is finite, the supremum and the essential supremum in (5.3) are identical. Set $\bar{\phi}_{t_i}(\xi) = \phi_{t_i}^{F_{t_i}}(1+\frac{\xi-1}{\sqrt{\Delta t_{i+1}}})\Delta t_{i+1}$. Let us prove that $\bar{\phi}_{t_i}$ is a dynamic penalty function. First of all note that, as by assumption $\sup_j \Delta t_j^N \leq 1$, the mapping $\xi \mapsto 1+\frac{\xi-1}{\sqrt{\Delta t_{i+1}}}$ from $D_{t_{i+1}}$ to $D_{t_{i+1}}$ is one-to-one. Hence,

$$\text{ess}\inf_{\xi\in D_{t_{i+1}}} \bar{\phi}_{t_i}(\xi) = \text{ess}\inf_{\xi\in D_{t_{i+1}}} \phi_{t_i}^{F_{t_i}}(\xi)\Delta t_{i+1} = 0.$$

Furthermore, clearly for $A \in \mathcal{F}_{t_i}^N$ and $\xi_1, \xi_2 \in D_{t_{i+1}}$

$$\begin{aligned}
\bar{\phi}_{t_i}(I_A\xi_1 + I_{A^c}\xi_2) &= \phi_{t_i}^{F_{t_i}}\Big(I_A\Big(1+\frac{\xi_1-1}{\sqrt{\Delta t_{i+1}}}\Big) + I_{A^c}\Big(1+\frac{\xi_2-1}{\sqrt{\Delta t_{i+1}}}\Big)\Big)\Delta t_{i+1} \\
&= \Big(I_A\phi_{t_i}^{F_{t_i}}(1+\frac{\xi_1-1}{\sqrt{\Delta t_{i+1}}}) + I_{A^c}\phi_{t_i}^{F_{t_i}}(1+\frac{\xi_2-1}{\sqrt{\Delta t_{i+1}}})\Big)\Delta t_{i+1} \\
&= I_A\bar{\phi}_{t_i}(\xi_1) + I_{A^c}\bar{\phi}(\xi_2).
\end{aligned}$$

Therefore, the $\bar{\phi}_{t_i}$ are indeed one-step penalty functions. Now the corollary follows from Proposition 4.5. $\square$

Note that another way of obtaining a DCRM from the generators F_{t_i} would be to glue them together in a time-consistent way on $t_0, \ldots, T$ by recursively defining

$$\rho_T(X) = -X \quad \text{and} \quad \rho_{t_i}(X) = F_{t_i}(-\rho_{t_{i+1}}(X)). \tag{5.4}$$

This procedure *always* leads to a DCRM (ρ_{t_i}) such that the restriction of ρ_{t_i} to $\mathcal{F}_{t_{i+1}}$ is equal to F_{t_i}. However, since by Proposition 4.5 this is equivalent to defining (ρ_{t_i}) by (4.1) with penalty functions $(\phi_{t_i}^{F_{t_i}})$, by the discussion above an additional scaling and tilting is needed (otherwise for instance all non-trivial coherent risk measures would blow up, see Proposition 5.1). Namely, there is an equivalent way to obtain the robust extension (5.3) by tilting the generators. Assume that we have generators F_{t_i} (given for instance by Examples 4.1). Define

$$\sigma_{t_i}(X) = \sqrt{\Delta t_{i+1}} F_{t_i}\left(\frac{1}{\sqrt{\Delta t_{i+1}}}\left(X - \mathbb{E}[X|\mathcal{F}_{t_i}^N]\right)\right).$$

Note that σ_{t_i} satisfies the axioms of a *general deviation measure* given in Rockafellar et al. (2006), except that sublinearity has to be replaced by convexity (in many of our examples, however, σ_{t_i} is actually a true general deviation measure in the sense of Rockafellar et al. (2006)). With a slight abuse of notation (which however is justified as we will see shortly) we *define* tilted one-period convex risk measures $\rho_{t_i,t_{i+1}}$ by

$$\rho_{t_i,t_{i+1}}(X) = \mathbb{E}[-X|\mathcal{F}_{t_i}^N] + \sqrt{\Delta t_{i+1}} \sigma_{t_i}(X) \tag{5.5}$$
$$= (1 - \sqrt{\Delta t_{i+1}})\mathbb{E}[-X|\mathcal{F}_{t_i}^N] + \sqrt{\Delta t_{i+1}} \hat{F}_{t_i}(X) \tag{5.6}$$

for any $\mathcal{F}_{t_{i+1}}^N$-measurable X, where we have set

$$\hat{F}_{t_i}(X) = \sqrt{\Delta t_{i+1}} F_{t_i}\left(\frac{X}{\sqrt{\Delta t_{i+1}}}\right), \quad t_i \in \{t_0, \ldots, t_{k-1}\}.$$

With these specific generators $\rho_{t_i,t_{i+1}}(X)$ we can define a DCRM ρ_t for $t \in \{t_0, \ldots, t_k\}$ by gluing the operators $\rho_{t_i,t_{i+1}}$ together, using (5.4) with F_{t_i} replaced by $\rho_{t_i,t_{i+1}}$, that is, by setting

$$\rho_T(X) = -X \quad \text{and} \quad \rho_{t_i}(X) = \rho_{t_i,t_{i+1}}(-\rho_{t_{i+1}}(X)). \tag{5.7}$$

Note that the restriction of (ρ_{t_i}) to $\mathcal{F}_{t_{i+1}}$ indeed is equal to $\rho_{t_i,t_{i+1}}$.

Proposition 5.4 *The DCRMs defined by (5.3) and by (5.5)-(5.7) coincide.*

Proof. We prove the assertion by showing that the operator ρ defined by (5.5)-(5.7) has indeed the dual representation given by (5.3). Since the σ-algebra $\mathcal{F}_T^N$ is finite, as it is generated by the Bernoulli random walk, we just need to look at finitely many atoms at the time instances t_i. Consider the functional $\mathbb{E}^{t_i}[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_{t_i}^N]$ from $L^0(\mathcal{F}_{t_{i+1}}^N)$ to $L^0(\mathcal{F}_{t_i}^N)$. Note that for the conjugate of $-\mathbb{E}^{t_i}$ we have on every atom of $\mathcal{F}_{t_i}^N$ that $\left(-\mathbb{E}^{t_i}\right)^*(\xi) = J_{\{\xi=1\}}$ where $J_{\{\xi=1\}}$ is the indicator function which is zero if $\xi = 1$ and infinity otherwise. Introduce the operation $(\phi_1 \square \phi_2)(\xi) = \inf_{\xi_1+\xi_2=\xi}\{\phi_1(\xi_1) + \phi_2(\xi_2)\}$. It is well-known that for dual conjugates of convex lower-semicontinuous functions the following relationships hold (see for instance Zalinescu (2002), Theorem 2.3.1):

- For every $\alpha > 0$, $\left(\alpha f(u)\right)^* = \alpha f^*(u/\alpha)$ and $\left(\alpha f(u/\alpha)\right)^* = \alpha f^*(u)$

- $(f + g)^*(u) = (f^* \Box g^*)(u).$

As the probability space is finite, ρ_{t_i} may be viewed as a real-valued convex function from a finite-dimensional Euclidean space to $\mathbb{R}$. It is well known that such functions are continuous in the Euclidean norm. Together with the definition of $\rho_{t_i, t_{i+1}}$ for $\xi_{t_{i+1}} \in D_{t_{i+1}}$ this yields on every atom of $\mathcal{F}_{t_i}^N$ that

$$
\begin{aligned}
\phi_{t_i}^{\rho_{t_i, t_{i+1}}}(\xi_{t_{i+1}}) &= \rho_{t_i, t_{i+1}}^*(\xi_{t_{i+1}}) \\
&= \Big(-(1 - \sqrt{\Delta t_{i+1}})\mathbb{E}^{t_i} + \sqrt{\Delta t_{i+1}}\hat{F}_{t_i} \Big)^*(\xi_{t_{i+1}}) \\
&= \Big(-(1 - \sqrt{\Delta t_{i+1}})\mathbb{E}^{t_i} \Big)^* \Box \Big(\sqrt{\Delta t_{i+1}}\hat{F}_{t_i} \Big)^*(\xi_{t_{i+1}}) \\
&= \inf_{\xi_1 + \xi_2 = \xi_{t_{i+1}}} \Big\{ J_{\{\xi_1/(1 - \sqrt{\Delta t_{i+1}}) = 1\}} + \Big(\sqrt{\Delta t_{i+1}}\hat{F}_{t_i} \Big)^*(\xi_2) \Big\} \\
&= \inf_{\xi_1 + \xi_2 = \xi_{t_{i+1}}} \Big\{ J_{\{\xi_1/(1 - \sqrt{\Delta t_{i+1}}) = 1\}} + \sqrt{\Delta t_{i+1}}\hat{F}_{t_i}^*\Big(\frac{\xi_2}{\sqrt{\Delta t_{i+1}}} \Big) \Big\} \\
&= \sqrt{\Delta t_{i+1}}\hat{F}_{t_i}^*\Big(\frac{\xi_{t_{i+1}} + \sqrt{\Delta t_{i+1}} - 1}{\sqrt{\Delta t_{i+1}}} \Big) \\
&= \Delta t_{i+1} F_{t_i}^*\Big(\frac{\xi_{t_{i+1}} + \sqrt{\Delta t_{i+1}} - 1}{\sqrt{\Delta t_{i+1}}} \Big) = \Delta t_{i+1}\phi_{t_i}^{F_{t_i}}\Big(1 + \frac{\xi_{t_{i+1}} - 1}{\sqrt{\Delta t_{i+1}}} \Big).
\end{aligned}
$$

As the upper continuity assumption of the first part of Proposition 4.5 is satisfied, we can conclude that indeed

$$
\rho_{t_i}(X) = \sup_{Q \in D} \mathbb{E}_Q\Big[-X - \sum_{j=i}^{k-1} \Delta t_{j+1}\phi_{t_j}^{F_{t_j}}\Big(1 + \frac{\xi_{t_{j+1}}^Q - 1}{\sqrt{\Delta t_{j+1}}} \Big) \Big| \mathcal{F}_{t_i}^N \Big].
$$

$\Box$

6 DCRMs and BSΔEs

In the sequel we will show that in the setting of Section 5 we can write the DCRM defined by (5.3) by means of a discrete BSDE (BSΔE). If $d = 1$ then by the predictable representation property of a one-dimensional Bernoulli random walk we have that for every $Z \in L^0(\mathcal{F}_T^N)$ there exists an $\mathcal{F}_{t_i}^N$-adapted process $(\gamma_{t_i})_{i \in \{0, \dots, k(N)-1\}}$ such that

$$
Z = \mathbb{E}[Z] + \sum_{i=0}^{k(N)-1} \gamma_{t_i} \Delta R_{t_{i+1}}^N.
$$

On the other hand if $d > 1$ by adding for every i, additional Bernoulli random variables $(\hat{B}_i^{N,l})_{l=1,\dots,2^d-d-1}$, such that for every fixed N and i, $\{(B_i^{N,l})_{l=1,\dots,d}, (\hat{B}_i^{N,l})_{l=1,\dots,2^d-d-1}\}$ are pairwise independent, we can define an auxiliary $\mathcal{F}^N$-adapted $(2^d\text{-}d\text{-}1)$-dimensional random walk

$$
\hat{R}_{t_{i+1}^N}^{N,l} = \sum_{j=1}^{i} \sqrt{\Delta t_j^N}\hat{B}_j^{N,l}, \quad i = 1, \dots, k(N), \; l = 1, \dots, 2^d - d - 1
$$

which is orthogonal to R^N, has pairwise independent components and independent increments, such that $(R^N, \hat{R}^N)$ have the predictable representation property. That is, for any $\mathcal{F}^N_{t_{i+1}}$-measurable random variable Y, there exist $\mathcal{F}^N_{t_i}$-measurable $\beta_{t_i}, \gamma_{t_i}, \hat{\gamma}_{t_i}$ such that

$$Y = \beta_{t_i} + \gamma_{t_i} \Delta R^N_{t_{i+1}} + \hat{\gamma}_{t_i} \Delta \hat{R}^N_{t_{i+1}},$$

see for instance Lemma 3.1 in Cheridito et al. (2009) or Lemma 4.3.1 and its discussion in Stadje (2009). Consequently, we can also find $\mathcal{F}^N_{t_i}$-measurable $\beta_{t_i}, \gamma_{t_i}$ and $\hat{\gamma}_{t_i}$ such that

$$\rho_{t_{i+1}}(X) = \beta_{t_i} + \gamma_{t_i} \Delta R^N_{t_{i+1}} + \hat{\gamma}_{t_i} \Delta \hat{R}^N_{t_{i+1}}.$$

This yields

$$
\begin{aligned}
\Delta \rho_{t_{i+1}}(X) &= \rho_{t_{i+1}}(X) - \rho_{t_i}(X) \\
&= \rho_{t_{i+1}}(X) - \rho_{t_i, t_{i+1}}(-\rho_{t_{i+1}}(X)) \\
&= \beta_{t_i} + \gamma_{t_i} \Delta R^N_{t_{i+1}} + \hat{\gamma}_{t_i} \Delta \hat{R}^N_{t_{i+1}} - \rho_{t_i, t_{i+1}}\left(-\beta_{t_i} - \gamma_{t_i} \Delta R^N_{t_{i+1}} - \hat{\gamma}_{t_i} \Delta \hat{R}^N_{t_{i+1}}\right) \\
&= -\rho_{t_i, t_{i+1}}\left(-\gamma_{t_i} \Delta R^N_{t_{i+1}} - \hat{\gamma}_{t_i} \Delta \hat{R}^N_{t_{i+1}}\right) + \gamma_{t_i} \Delta R^N_{t_{i+1}} + \hat{\gamma}_{t_i} \Delta \hat{R}^N_{t_{i+1}} \quad (6.1)
\end{aligned}
$$

where we have used cash invariance in the last equation. From now on we index again everything by N. For $z_1 \in \mathbb{R}^d$ and $z_2 \in \mathbb{R}^{2^d - d - 1}$ let

$$
\begin{aligned}
g^N(t_i^N, z_1, z_2) &= \frac{1}{\Delta t^N_{i+1}} \rho^N_{t_i^N, t_{i+1}^N}\left(-z_1 \Delta R^N_{t_{i+1}^N} - z_2 \Delta \hat{R}^N_{t_{i+1}^N}\right) \\
&= F^N_{t_i^N}\left(-\frac{1}{\sqrt{\Delta t^N_{i+1}}}\left(\sqrt{\Delta t^N_{i+1}}\left(z_1 B^N_{i+1} + z_2 \hat{B}^N_{i+1}\right)\right)\right) \\
&= F^N_{t_i^N}\left(-z_1 B^N_{i+1} - z_2 \hat{B}^N_{i+1}\right) \quad (6.2)
\end{aligned}
$$

where we have used (5.6). Recall that the $B^N_{i+1}, \hat{B}^N_{i+1}$ are the Bernoulli variables which were introduced to generate the random walks $R^N, \hat{R}^N$. From (6.1) we get

$$\Delta \rho^N_{t_{i+1}^N}(X) = -\Delta t^N_{i+1} g^N(t_i^N, \gamma^N_{t_i^N}, \hat{\gamma}^N_{t_i^N}) + \gamma^N_{t_i^N} \Delta R^N_{t_{i+1}^N} + \hat{\gamma}^N_{t_i^N} \Delta \hat{R}^N_{t_{i+1}^N}.$$

This entails

$$
\begin{aligned}
\rho^N_T(X) - \rho^N_{t_i^N}(X) &= \sum_{j=i}^{k(N)-1} \Delta \rho^N_{t_{j+1}^N}(X) \\
&= -\sum_{j=i}^{k(N)-1} g^N(t_j^N, \gamma^N_{t_j^N}, \hat{\gamma}^N_{t_j^N}) \Delta t^N_{j+1} + \sum_{j=i}^{k(N)-1}\left(\gamma^N_{t_j^N} \Delta R^N_{t_{j+1}^N} + \hat{\gamma}^N_{t_j^N} \Delta \hat{R}^N_{t_{j+1}^N}\right).
\end{aligned}
$$

From $\rho^N_T(X^N) = -X^N$ we obtain

$$\rho^N_{t_i^N}(X^N) = -X^N + \sum_{j=i}^{k(N)-1} g^N(t_j^N, \gamma^N_{t_j^N}, \hat{\gamma}^N_{t_j^N}) \Delta t^N_{j+1} - \sum_{j=i}^{k(N)-1}\left(\gamma^N_{t_j^N} \Delta R^N_{t_{j+1}^N} + \hat{\gamma}^N_{t_j^N} \Delta \hat{R}^N_{t_{j+1}^N}\right).$$

16

Setting $\gamma_s^N = \gamma_{t_i^N}^N$ and $\hat{\gamma}_s^N = \hat{\gamma}_{t_i^N}^N$ for $t_i^N \leq s < t_{i+1}^N$ we can write the last equation also in the form of a discrete backward stochastic differential equation (BSΔE)

$$\rho_t^N(X^N) = -X^N + \int_t^T g^N(s-, \gamma_{s-}^N, \hat{\gamma}_{s-}^N) ds^N - \int_t^T \gamma_{s-}^N dR_s^N - \int_t^T \hat{\gamma}_{s-}^N d\hat{R}_s^N. \qquad (6.3)$$

Thus we have proved the following proposition.

Proposition 6.1 *The robust extension ρ^N defined by (5.3) is the solution of the BSΔE (6.3).*

Remark 6.2 While Definition 5.2 requires the convexity of the generators, to arrive at (6.3) from (5.5)-(5.7) only the cash invariance of $F_{t_i^N}^N$ is needed. Thus, defining a robust extension by (5.5)-(5.7) for one-period operators satisfying cash invariance (but possibly not convexity) we can also obtain (6.3). In particular, also $V@R$ can be extended to discrete time in this way and the extension satisfies (6.3).

The introduction of the $\hat{\gamma}_{t_i^N}^N$ is due to the fact that the Bernoulli random walk in higher dimensions does not have the predictable representation property. However, when N gets large the random walk converges to a Brownian motion W which does have the predictable representation property. Thus, we might expect that for large N the $\hat{\gamma}^N$ converge to zero.

From now on we assume

(B1): There exists a function $g(t, \omega, z_1)$ satisfying the assumptions stated in (H1) or (H2) such that for every $z_1 \in \mathbb{R}^d$

$$\sup_{0 \leq t \leq T} |g^N(t, z_1, 0) - g(t, z_1)| \stackrel{N \to \infty}{\to} 0 \ \ \text{in } L^2.$$

Note that (B1) is satisfied if g^N is independent of ω, t and of N. This is the case the following condition holds:

(B2) The homogenous case: $F_{t_i^N}^N(-z_1 B_{i+1}^N - z_2 \hat{B}_{i+1}^N)$ is deterministic and independent of N and i.

Assumption (B2), and thus also (B1), is satisfied in the Examples 4.1. Notice that $F_{t_i^N}^N(-z_1 B_{i+1}^N - z_2 \hat{B}_{i+1}^N)$ is always deterministic provided $F_{t_i}^N$ is law-invariant under $P[\cdot|\mathcal{F}_{t_i}^N]$.

The BSDE corresponding to (6.3) should be

$$\rho_t(X) = -X + \int_t^T g(s, Z_s) ds - \int_t^T Z_s dW_s, \ \ t \in [0, T]. \qquad (6.4)$$

Now, we want that if X^N converges to some random variable X, then in some sense of process convergence, ρ^N tends to the process ρ which appears in the solution of the corresponding BSDE (6.4).

Proposition 6.3 *Suppose that the the operators $F_{t_i^N}^N$ are monotone and cash invariant (not necessarily convex). Then for every t_i^N, $g^N(t_i^N, z_1, z_2)$, defined by (6.2), is uniformly Lipschitz continuous in (z_1, z_2) with Lipschitz constant $\max(\sqrt{d}, \sqrt{2^d - d - 1})$.*

Proof. Let $z_1, z_1' \in \mathbb{R}^d$ and $z_2, z_2' \in \mathbb{R}^{2^d - d - 1}$. From (6.2),

$$
\begin{aligned}
g^N(t_i^N, z_1, z_2) &= F_{t_i^N}^N(-z_1 B_{i+1}^N - z_2 \hat{B}_{i+1}^N) \\
&\leq F_{t_i^N}^N\big(- z_1' B_{i+1}^N - z_2' \hat{B}_{i+1}^N - ||(-z_1 + z_1')B_{i+1}^N + (-z_2 + z_2')\hat{B}_{i+1}^N||_\infty \big) \\
&= F_{t_i^N}^N\big(- z_1' B_{i+1}^N - z_2' \hat{B}_{i+1}^N \big) + ||(-z_1 + z_1')B_{i+1}^N + (-z_2 + z_2')\hat{B}_{i+1}^N||_\infty \\
&\leq F_{t_i^N}^N\big(- z_1' B_{i+1}^N - z_2' \hat{B}_{i+1}^N \big) + ||(-z_1 + z_1')B_{i+1}^N||_\infty + ||(-z_2 + z_2')\hat{B}_{i+1}^N||_\infty \\
&\leq g^N(t_i^N, z_1', z_2') + \sqrt{d}| - z_1 + z_1'| + \sqrt{2^d - d - 1}| - z_2 + z_2'|.
\end{aligned}
$$

We have used the monotonicity in the first inequality and cash invariance in the second equality. In the last inequality we have applied Cauchy's inequality. $\square$

Now everything is ready for the following convergence theorem.

Theorem 6.4 *Assume that (B1) holds and the $\mathcal{F}_T^N$-measurable discrete-time payoffs X^N converge to the $\mathcal{F}_T$-measurable continuous-time payoff X in L^2. Let $\rho_t^N(X^N)$ be the robust extension of given generators $(F_{t_i^N}^N)_{i=0,\ldots,k-1}$ and the dynamic continuous time risk measure $\rho_t(X)$ be defined by (6.4). Then we have*

$$
\sup_{0 \leq t \leq T} |\rho_t^N(X^N) - \rho_t(X)| \stackrel{N\to\infty}{\to} 0 \quad in \ L^2. \tag{6.5}
$$

Furthermore,

$$
\mathbb{E}\Big[\int_0^T \big(|\gamma_{s-}^N - \gamma_{s-}|^2 + |\hat{\gamma}_{s-}^N|^2 \big) ds \Big] \to 0 \ as \ N \to \infty. \tag{6.6}
$$

Proof. By Proposition 6.1 ρ^N satisfies the BSΔE (6.3). Because of Proposition 6.3 the driver g^N is uniformly Lipschitz, with Lipschitz constant independent of N. Furthermore, $g^N(t, 0, 0) = 0$ for all $N \in \mathbb{N}$. This together with (B1) yields that the assumptions of Theorem 12 in Briand et al. (2002) are satisfied and we can conclude that (6.5)-(6.6) hold. $\square$

Remark 6.5 Actually for $d > 1$ a slight generalization of the result by Briand et al. (2002) is needed because in their paper drivers of the BSΔEs, which are independent of y, have the form $g^N(s-, \omega, \gamma_{s-}^N)$ while in our case we consider drivers $g^N(s-, \omega, \gamma_{s-}^N, \hat{\gamma}_{s-}^N)$. However, in the multi-dimensional situation the proof in Briand et al. (2002) can easily be extended to this case.

Thus, the continuous-time convex risk measure ρ satisfying a BSDE with Lipschitz continuous driver g can be interpreted very naturally as the limit of discrete-time risk measures ρ^N. Now, how can one find $\mathcal{F}_T^N$-measurable X^N converging to a given X? Two possibilities are of particular interest:

   (a) Let $X \in L^2(\mathcal{F}_T)$ be of the form $X = h(W_T)$ where $h : \mathbb{R} \to \mathbb{R}$ is a continuous function which grows at most polynomially. Then we can define $X^N = h(R_T^N)$.

   (b) For general $X \in L^2(\mathcal{F}_T)$ define $X^N = \mathbb{E}[X|\mathcal{F}_T^N]$.

   In both cases we have

$$
\sup_{0 \leq t \leq T} |\rho_t^N(X^N) - \rho_t(X)| \to 0
$$

in L^2. Theorem 6.4 also shows that in a certain sense the risk modeling in discrete time with the tilted operators ρ^N is robust. Let X^N be a L^2 Cauchy sequence of discrete-time terminal conditions. Then we have seen that for every $\varepsilon > 0$ there exists an N_0 such that for all $N, M \geq N_0$ we have $|\rho_0^N(X^N) - \rho_0^M(X^M)| \leq \varepsilon$. Thus, refining the time grid only leads to small changes in the risk evaluation from a certain index on. Contrary to that, Proposition 5.1 shows that if the discrete-time risk measure is constructed as in (5.2) without further scaling, then, if for instance the one-period risk measures are coherent, the discrete-time risk measurement will blow up when more and more time instances are taken into account. So in particular for Value at Risk, Average Value at Risk or the Semi-deviation our additional scaling is necessary. For Average Value at Risk it has been observed before that (5.2) leads to a too conservative risk measurement, see Roorda and Schumacher (2007).

7 Examples of one-period convex risk measures extended to continuous time

7.1 Semi-Deviation

Suppose that the generators $F_{t_i^N}^N$ correspond to the semi-deviation risk measure from Example 4.1. We get from (6.2) for $t_i^N \leq t < t_{i+1}^N$, $z_1 \in \mathbb{R}^d$ and $z_2 \in \mathbb{R}^{2^d - d - 1}$

$$g^N(t, z_1, z_2) = \lambda \left| \mathbb{E}\left[\left(-\sum_{l=1}^d z_1^l B_{i+1}^{N,l} - \sum_{l=1}^{2^d - d - 1} z_2^l \hat{B}_{i+1}^{N,l}\right)_-^q\right]\right|^{1/q}.$$

These driver functions are in fact independent of N and t. In particular we may write $g(z^1, z^2)$. By Proposition 6.1 the robust extension ρ^N of the semi-deviation one-period risk measures to discrete time is given by

$$\rho_t^N(X^N) = -X^N + \int_t^T g(\gamma_{s-}^N, \hat{\gamma}_{s-}^N)ds^N - \int_t^T \gamma_{s-}^N dR_s^N - \int_t^T \hat{\gamma}_{s-}^N d\hat{R}_s^N.$$

The robust extension (which corresponds to using locally tilted semi-deviations) should be used if information is coming in frequently to ensure that updating does not lead the risk measure to blow up. For $z \in \mathbb{R}^d$ with $z = (z^1, \ldots, z^d)$ define

$$g(z) = g^N(z, 0) = \lambda \left| \frac{1}{2^d} \sum_{(k_1, \ldots, k_d) \in \{-1, 1\}^d} \left(-\sum_{l=1}^d k_l z^l\right)_-^q \right|^{1/q}.$$

Let $(\rho_t(X), Z_t)$ be the solution of

$$\rho_t(X) = -X + \int_t^T g(Z_s)ds - \int_t^T Z_s dW_s.$$

Then Theorem 6.4 yields that for every $X \in L^2(\mathcal{F}_T)$ and every sequence $X^N \in L^0(\mathcal{F}_T^N)$ converging to X in L^2, we have

$$\sup_{0 \leq t \leq T} |\rho_t^N(X^N) - \rho_t(X)| \overset{N \to \infty}{\Rightarrow} 0 \text{ in } L^2.$$

7.2 Value at Risk

For the generators $F_{t_i^N}^N$ being equal to Value at Risk, we obtain from (6.2) for $t_i^N \leq t < t_{i+1}^N$, $z_1 \in \mathbb{R}^d$ and $z_2 \in \mathbb{R}^{2^d-d-1}$

$$g^N(t, z_1, z_2) = V@R_{t_i^N}^{N,\alpha}\left(-\sum_{l=1}^{d} z_1^l B_{i+1}^{N,l} - \sum_{l=1}^{2^d-d-1} z_2^l \hat{B}_{i+1}^{N,l} \right).$$

As $g^N(t, z_1, z_2)$ is independent of N and t we may write $g(z^1, z^2)$. Define

$$x_p = p\text{-th largest element of the set}$$
$$\{(-1)^{k_1} z^1 + \ldots (-1)^{k_d} z^d | k_l \in \{1, 2\}, l = 1, \ldots, d\} \tag{7.1}$$

for $p = 1, \ldots, 2^d$. Note that we have $x_p = -x_{2^d+1-p}$. Now extending Value at Risk to discrete time in the way proposed in Remark 6.2 gives

$$\rho_t^N(X^N) = -X^N + \int_t^T g(\gamma_{s-}^N, \hat{\gamma}_{s-}^N) ds^N - \int_t^T \gamma_{s-}^N dR_s^N - \int_t^T \hat{\gamma}_{s-}^N d\hat{R}_s^N. \tag{7.2}$$

For $z \in \mathbb{R}^d$ let

$$g^\alpha(z) = g(z, 0) = -x_{2^d - \lfloor \alpha 2^d \rfloor}. \tag{7.3}$$

For example if $\alpha < \dfrac{1}{2^d}$,

$$g^\alpha(z) = -x_{2^d} = |z^1| + \ldots + |z^d|.$$

If $\dfrac{1}{2^d} \leq \alpha < \dfrac{2}{2^d}$,

$$g^\alpha(z) = w^{[2]} + \ldots + w^{[d]} - w^{[1]}$$

where we applied the order statistic $[\cdot]$ to the components of the d-dimensional vector $w = (|z^1|, \ldots, |z^d|)$. Note that g^α is not convex if $1 > \alpha \geq 1/2^d$, which is due to the lack of convexity of V@R. Let $(\rho_t(X), Z_t)$ be the solution of

$$\rho_t(X) = -X + \int_t^T g^\alpha(Z_s) ds - \int_t^T Z_s dW_s.$$

Since $(\rho_{t_i^N}^N)_{i=0,\ldots,k}$ is defined directly by (7.2) instead of (5.3) (because of the non-convexity of V@R), we can not use Theorem 6.4. However, we can apply Theorem 12 from Briand et al. (2002) directly. This yields that for every $X \in L^2(\mathcal{F}_T)$ and every sequence X^N converging to X in L^2 we have

$$\sup_{0 \leq t \leq T} |\rho_t^N(X^N) - \rho_t(X)| \overset{N \to \infty}{\to} 0 \text{ in } L^2.$$

7.3 Average Value at Risk

From (6.2) we have for $t_i^N \leq t < t_{i+1}^N$, $z_1 \in \mathbb{R}^d$ and $z_2 \in \mathbb{R}^{2^d-d-1}$

$$g^N(t, z_1, z_2) = AV@R_{t_i^N}^{N,\lambda}\Big(-\sum_{l=1}^d z_1^l B_{i+1}^{N,l} - \sum_{l=1}^{2^d-d-1} z_2^l \hat{B}_{i+1}^{N,l}\Big).$$

As g^N is independent of N and t, we get the continuous-time driver, for $z \in \mathbb{R}^d$,

$$g(z) = \frac{1}{\alpha}\int_0^\alpha g^\lambda(z)d\lambda = -\frac{1}{\alpha}\Big(x_{2^d-\lceil 2^d\alpha\rceil+1}\Big(\alpha - \frac{\lceil 2^d\alpha\rceil-1}{2^d}\Big) + \frac{1}{2^d}\sum_{j=1}^{\lceil 2^d\alpha\rceil-1} x_{2^d-j+1}(z)\Big),$$

where g^λ and $(x_p)_{p=1,\ldots,2^d}$ were defined in (7.3) and (7.1). Let ρ^N be the robust extension of Average Value at Risk and let $(\rho_t(X), Z_t)$ be the solution of

$$\rho_t(X) = -X + \int_t^T g(Z_s)ds - \int_t^T Z_s dW_s.$$

Theorem 6.4 yields that for every $X \in L^2(\mathcal{F}_T)$ and every discrete time sequence X^N converging to X in L^2, we have

$$\sup_{0\leq t\leq T} |\rho_t^N(X^N) - \rho_t(X)| \stackrel{N\to\infty}{\to} 0 \text{ in } L^2.$$

7.4 The Gini risk measure

Let $(\rho_{t_i^N}^N)_{i=0,\ldots k(N)}$ be the robust extension of the one-period Gini risk measures $V_{t_i^N}^{N,\theta}$ from Example 4.1. Let $z_1 \in \mathbb{R}^d$ and $z_2 \in \mathbb{R}^{2^d-d-1}$. From (6.2) and the definition of the Gini risk measure we have for $t_i^N \leq t < t_{i+1}^N$,

$$g^N(t, z_1, z_2) = V_{t_i^N}^{N,\theta}\Big(-\sum_{l=1}^d z_1^l B_{i+1}^{N,l} - \sum_{l=1}^{2^d-d-1} z_2^l \hat{B}_{i+1}^{N,l}\Big).$$

As g^N is independent of t and N we can set

$$g(z) = g^N(t, z, 0) = \sup_q \Big\{-q^t x - \frac{1}{2\theta}(2^d\|q\|_2^2 - 1)\Big\}, \tag{7.4}$$

where $(x_p)_{p=1,\ldots,2^d}$ was defined in (7.1) and the supremum is taken over all $q = (q_1,\ldots,q_{2^d}) \in \mathbb{R}^{2^d}$ with $0 \leq q_p$ and $\sum_{p=1}^{2^d} q_p = 1$.

Let us derive an explicit formula for the driver in continuous time. Since (7.4) is a concave optimization problem and Slater's condition is satisfied, the solution is uniquely determined by the Karush-Kuhn-Tucker conditions:

$$q_p \geq 0,\ \lambda_p \geq 0,\ \lambda_p q_p = 0,\ x_p + \frac{2^d q_p}{\theta} - \nu - \lambda_p = 0 \text{ for } p = 1,\ldots,2^d \text{ and } 1^t q = 1.$$

From the last equation, $x_p + \frac{2^d q_p}{\theta} - \nu = \lambda_p$. Thus, we have to solve the system

$$\frac{\theta}{2^d}(\nu - x_p) \leq q_p, \tag{7.5}$$

21

$$(x_p + \frac{2^d q_p}{\theta} - \nu)q_p = 0,$$

$$q_p \geq 0 \text{ for } p = 1, \ldots, 2^d \text{ and } 1^t q = 1.$$

Therefore, we can conclude that $q_p = 0$ or $\nu = x_p + \frac{2^d q_p}{\theta}$ for every $p = 1, \ldots, 2^d$. This and (7.5) yields,

$$q_p = \frac{\theta}{2^d} \max(0, \nu - x_p) \text{ for } p = 1, \ldots, 2^d. \tag{7.6}$$

Consequently, ν is obtained as the unique solution of the equation

$$\frac{\theta}{2^d} \sum_{p=1}^{2^d} \max(0, \nu - x_p) = 1. \tag{7.7}$$

Let us look at (7.6)-(7.7) a little closer. Note that $\nu - x_p$ is increasing in p. Therefore, if $\nu - x_{w-1} \leq 0$, for an index $w - 1$, we have $\nu - x_p \leq 0$ for all $p \leq w - 1$. Thus, there exists an index $w \in \{1, \ldots, 2^d\}$ such that $q_p = (\theta/2^d)(\nu - x_p)$ for $p = 2^d, \ldots, w$ and $q_p = 0$ for $p = 1, \ldots, w - 1$. Hence, $\sum_{p=w}^{2^d}(\nu - x_p) = 2^d/\theta$ and $\nu - x_{w-1} \leq 0$. This yields

$$\nu = \frac{2^d}{\theta(2^d + 1 - w)} + \frac{\sum_{p=w}^{2^d} x_p}{2^d + 1 - w}$$

and w is uniquely determined as follows:

$$w = \sup \left\{ w \Big| \text{for all } j \in \{2^d, \ldots, w\} : \quad \frac{2^d}{\theta(2^d + 1 - w)} > -\frac{\sum_{p=j}^{2^d} x_p}{2^d + 1 - w} + x_j \right.$$

$$\left. \text{and } \frac{2^d}{\theta(2^d + 1 - w)} \leq -\frac{\sum_{p=w}^{2^d} x_p}{2^d + 1 - w} + x_{w-1} \right\} \vee 1. \tag{7.8}$$

Since x_p and therefore also w depend on z, we will subsequently write $w(z)$ and $x_p(z)$. Inserting the values for ν and q_p, $p = 1, \ldots, 2^d$, into (7.4) we get

$$g(z) = -\frac{\theta}{2^d} \sum_{p=w(z)}^{2^d} \left[\nu x_p(z) - x_p^2(z) \right]$$

$$- \frac{2^{d-1}}{\theta} \sum_{p=w(z)}^{2^d} \left[(\theta/2^d)^2(\nu^2 - 2\nu x_p(z) + x_p^2(z)) \right] + \frac{1}{2\theta}$$

$$= -\frac{\theta}{2^d} \sum_{p=w(z)}^{2^d} \left[\nu x_p(z) - x_p^2(z) \right] - \frac{\theta}{2^{d+1}} \sum_{p=w(z)}^{2^d} \left[(\nu^2 - 2\nu x_p(z) + x_p^2(z)) \right] + \frac{1}{2\theta}$$

$$= \frac{\theta}{2^{d+1}} \left(-(2^d + 1 - w(z))\nu^2 + \sum_{p=w(z)}^{2^d} x_p^2(z) \right) + \frac{1}{2\theta}.$$

Hence,

$$g(z) = \frac{\theta}{2^{d+1}} \left(-\frac{2^{2d}}{\theta^2(2^d + 1 - w(z))} - 2^{d+1} \frac{\sum_{p=w(z)}^{2^d} x_p}{\theta(2^d + 1 - w(z))} - \frac{\left(\sum_{p=w(z)}^{2^d} x_p(z)\right)^2}{2^d + 1 - w(z)} \right.$$

$$\left. + \sum_{p=w(z)}^{2^d} x_p^2(z) \right) + \frac{1}{2\theta}.$$

Thus, the driver for the BSDE extension of the Gini risk measure is given by

$$g(z) = -\frac{2^{d-1}}{\theta(2^d + 1 - w(z))} + \frac{1}{2\theta} - \frac{\sum_{j=w(z)}^{2^d} x_i(z)}{2^d + 1 - w(z)}$$

$$+ \frac{\theta}{2^{d+1}} \left(\sum_{j=w(z)}^{2^d} x_j^2(z) - \frac{\left(\sum_{j=w(z)}^{2^d} x_i(z)\right)^2}{2^d + 1 - w(z)} \right). \tag{7.9}$$

Note that if $z = 0$ then $w = 1$ and hence by the symmetry of the x_i we have $g(0) = 0$ which is necessary for the risk measure to be normalized. In the special case $d = 1$ we get from (7.8) that $w = 1$ if $|z| < 1/\theta$, and $w = 2$ if $|z| \geq 1/\theta$. Thus, (7.9) implies that g is equal to the Huber penalty function

$$g(z) = \begin{cases} |z| - \dfrac{1}{2\theta}, & \text{if } |z| \geq 1/\theta \\ \dfrac{\theta}{2} z^2, & \text{if } |z| < 1/\theta. \end{cases}$$

The Huber penalty function plays an important role in regression. It penalizes large errors with the L^1 norm (for robustness reasons) and small errors with the L^2 norm. It is continuously differentiable. Now let $(\rho_t(X), Z_t)$ be the solution of

$$\rho_t(X) = -X + \int_t^T g(Z_s)ds - \int_t^T Z_s dW_s.$$

Then Theorem 6.4 yields

$$\sup_{0 \leq t \leq T} |\rho_t^N(X^N) - \rho_t(X)| \overset{N \to \infty}{\to} 0 \text{ in } L^2$$

for every $X \in L^2(\mathcal{F}_T)$ and every discrete time sequence X^N converging to X in L^2.

8 Appendix

Proof of Proposition 5.1. Define $g(z_1, z_2) = F_{t_i}^N(z_1 B_{i+1}^N + z_2 \hat{B}_{i+1}^N)$. As the right hand side is independent of N and i, g is well-defined. Following the lines of the proof of Proposition 6.1, with the tilted generator $\rho_{t_i, t_{i+1}}$ replaced by $F_{t_i}^N$, we can see, using the assumption that $F_{t_i}^N$ are coherent, that for every discrete time payoff X^N, ρ^N constructed by (5.2) satisfies the BSΔE

$$\rho_{t_i^N}^N(X^N) = -X^N + \sum_{j=i}^{k(N)-1} (\Delta t_{j+1}^N)^{-1/2} g(\gamma_{t_j^N}^N, \hat{\gamma}_{t_j^N}^N) \Delta t_{j+1}^N - \sum_{j=i}^{k(N)-1} \left(\gamma_{t_j^N}^N \Delta R_{t_{j+1}^N}^N + \hat{\gamma}_{t_j^N}^N \Delta \hat{R}_{t_{j+1}^N}^N \right). \tag{8.1}$$

Let us introduce suitable X^N's. Define

$$X^N = -z_0 R_T^N \text{ and } X = -z_0 W_T.$$

Clearly, X^N converges in L^2 to X. Moreover, using the tools of weak convergence of filtrations, Proposition 2 and the second point of Remark 1 in Coquet et al. (2001) yield that $\mathbb{E}[X^N|\mathcal{F}_\cdot^N]$ converges to $\mathbb{E}[X|\mathcal{F}_\cdot]$ uniformly in probability. Passing to a subsequence if necessary we may assume that $\mathbb{E}[X^N|\mathcal{F}_\cdot^N]$ converges to $\mathbb{E}[X|\mathcal{F}_\cdot]$ uniformly a.s.

Now let us prove that $\rho_t^N(X^N) \stackrel{N\to\infty}{\to} \infty$ for $t \in [0,T)$. The process Y^N defined by

$$Y_{t_i^N}^N = \sum_{j=i}^{k(N)-1} g(z_0,0)\sqrt{\Delta t_{j+1}^N} - \mathbb{E}[X^N|\mathcal{F}_{t_i^N}^N]$$

and $Y_t^N = Y_{t_i^N}^N$ for $t_i^N \leq t < t_{i+1}^N$ is $\mathcal{F}^N$-adapted, and

$$Y_{t_i^N}^N = \sum_{j=i}^{k(N)-1} (\Delta t_{j+1}^N)^{-1/2} g(z_0,0) \Delta t_{j+1}^N + \sum_{j=0}^{i-1} z_0 \Delta R_{t_{j+1}^N}^N$$

$$= -X^N + \sum_{j=i}^{k(N)-1} (\Delta t_{j+1}^N)^{-1/2} g(z_0,0) \Delta t_{j+1}^N - \sum_{j=i}^{k(N)-1} z_0 \Delta R_{t_{j+1}^N}^N,$$

where we used in the last equation that $X^N = -\sum_{j=0}^{k(N)-1} z_0 \Delta R_{t_{j+1}^N}^N$. Hence, Y^N is a solution of the BSΔE (8.1) with terminal condition $-X^N$ and $\gamma_{t_j}^N = z_0$ and $\hat{\gamma}_{t_j}^N = 0$. Therefore, $\rho_t^N(X^N) = Y_t^N$. Now let $\Pi^N = \max_{j=0,\dots,k(N)-1} \Delta t_{j+1}^N$. Since $g(z_0,0) > 0$ we get for every $t \in [0,T)$

$$\rho_t^N(X^N) = Y_t^N = -\mathbb{E}[X^N|\mathcal{F}_t^N] + \sum_{j:\, t_{j+1}^N > t} (\Delta t_{j+1}^N)^{-1/2} g(z_0,0) \Delta t_{j+1}^N$$

$$\geq -\mathbb{E}[X^N|\mathcal{F}_t] + (\Pi^N)^{-1/2} g(z_0,0) \sum_{j:\, t_{j+1}^N > t} \Delta t_{j+1}^N$$

$$\geq -\mathbb{E}[X^N|\mathcal{F}_t^N] + (\Pi^N)^{-1/2} g(z_0,0)(T - t - \Pi^N).$$

As for every fixed t, $\mathbb{E}[X^N|\mathcal{F}_t^N]$ converges a.s. to the finite random variable $\mathbb{E}[X|\mathcal{F}_t]$, and $\Pi^N \to 0$, it follows that $\liminf_N \rho_t^N(X^N) = \infty$. $\square$

References

[1] Artzner, P., Delbaen, F., Eber, J.M., Heath, D., (1997). Thinking coherently. *RISK* 10, 68-71.

[2] Artzner, P., Delbaen, F., Eber, J.M., Heath, D., (1999). Coherent measures of risk. *Mathematical Finance* 9, 203-228.

[3] Artzner, P., Delbaen, F., Eber, J.M., Heath, D., Ku, H., (2007). Coherent multiperiod risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152, 5-22.

[4] Barrieu, P., El Karoui, N., (2004). Optimal derivative design under dynamic risk measures. In:*Mathematics of Finance, Contemporary Mathematics* 351, 13-26.

[5] Barrieu, P., El Karoui, N., (2009). Pricing, hedging and optimally designing derivatives via minimization of risk measures. In: R. Carmona (ed.) *Volume on Indifference Pricing*, Princeton University Press.

[6] Bion-Nadal, J., (2008). Dynamic risk measures: time consistency and risk measures from BMO martingales. *Finance and Stochastic* 12, 219-244.

[7] Briand, P., Delyon, B., Mémin, J., (2002). On the robustness of backward stochastic differential equations. *Stochastic Processes and their Applications* 97, 229-253.

[8] Chen, Z., Epstein, L.G., (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70, 1403-1443.

[9] Cheridito, P., Delbaen, F., Kupper, M., (2006). Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probabability* 11, 57-106.

[10] Cheridito, P., Kupper, M., (2006). Composition of time-consistent dynamic monetary risk measures in discrete time. Preprint.

[11] Cheridito, P., Horst, U., Kupper, M., Pirvu, T.A., (2009). Equilibrium pricing in incomplete markets under translation invariant preferences. Preprint.

[12] Coquet, F.,Hu, Y., Mémin J., Peng, S., (2002). Filtration-consistent nonlinear expectations and related g-expectations. *Probability Theory and Related Fields* 123, 1-27.

[13] Coquet, F., Mémin, J., Slominski, J., (2001). On weak convergence of filtrations: *Séminaire de Probabilités XXXV, Lecture Notes in Mathematics.* Springer.

[14] Delbaen, F., (2002). Coherent risk measures on general probability spaces. In: P. J. Schönbucher and K. Sandmann (eds.) *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann*, Springer.

[15] Delbaen, F., (2006). The Structure of m-Stable Sets and in Particular of the Set of Risk Neutral Measures. *Lecture Notes in Mathematics 1874. In Memoriam Paul-André Meyer*, 215-258, Springer.

[16] Deprez, O., Gerber, H.U., (1985). On convex principles of premium calculation. *Insurance: Mathematics and Economics* 4, 179-189.

[17] Detlefsen, K., Scandolo, G., (2005). Conditional and dynamic convex risk measures. *Finance and Stochastic* 9, 539-561.

[18] Duffie, D., Epstein, L.G., (1992). Stochastic differential utility. *Econometrica* 60, 353-394.

[19] Epstein, L.G., Schneider, M., (2003). Recursive multiple-priors. *Journal of Economic Theory* 113, 1-31.

[20] Föllmer, H., Schied, A., (2002). Convex measures of risk and trading constraints. *Finance and Stochastic* 6, 429-447.

[21] Föllmer, H., Schied, A., (2004). *Stochastic Finance: An Introduction in Discrete Time*, 2nd ed. De Gruyter, Berlin.

[22] Föllmer, H., Penner, I., (2006). Convex risk measures and the dynamics of their penalty functions. *Statistics and Decisions* 24, 61-96.

[23] Frei, C., Malamud, S., Schweizer, M., (2009). Convexity bounds for BSDE solutions with applications to indifference valuation. Preprint.

[24] Frittelli, M., Rosazza Gianin, E., (2002). Putting order in risk measures. *Journal of Banking and Finance* 26, 1473-1486.

[25] Frittelli, M., Rosazza Gianin, E., (2004). Dynamic convex risk measures. In: G. Szeg (ed.) *Risk Measures for the 21st Century*, Chapter 12. Wiley Finance.

[26] Goovaerts, M.J., Kaas, R., Dhaene, J., Tang, Q, (2003). A unified approach to generate risk measures. *Astin Bulletin* 33, 173-191.

[27] Goovaerts, M.J., Kaas, R., Dhaene, J., Tang, Q, (2004a). Some new classes of consistent risk measures. *Insurance: Mathematics and Economics* 34, 505-516.

[28] Goovaerts, M.J., Kaas, R., Laeven, R.J.A., Tang, Q, (2004b). A comonotonic image of independence for additive risk measures. *Insurance: Mathematics and Economics* 35, 581-594.

[29] Hu, Y., Ma, J., Peng, S., Yao, S., (2008). Representation theorems for quadratic F-consistent nonlinear expectations. *Stochastic Processes and their Applications* 118, 1518-1551.

[30] Ikeda, N., Watanabe, S., (1989). *Stochastic Differential Equations and Diffusion Processes*. Amsterdam, North-Holland.

[31] Jiang, L., (2008). Convexity, translation invariance and subadditivity for g-expectations and related risk measures. *Annals of Applied Probabability* 18, 245-258.

[32] Jobert, A., Rogers, L.C.G., (2008). Valuations and dynamic convex risk measures. *Mathematical Finance* 18, 1-22.

[33] Kazamaki, N., (1994). *Continuous Exponential Martingales and BMO.* Lecture Notes in Mathematics 1579, Springer.

[34] Klöppel, S., Schweizer, M., (2007). Dynamic indifference valuation via convex risk measures. *Mathematical Finance* 17, 599-627.

[35] Kobylanski, M., (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability* 28, 558-602.

[36] Landsman, Z., Sherris, M., (2001). Risk measures and insurance premium principles. *Insurance: Mathematics and Economics* 29, 103-115.

[37] Maccheroni, F., Marinacci, M., Rustichini, A., Taboga, M., (2004). Portfolio Selection with Monotone Mean-Variance Preferences. ICER Working Papers, *Applied Mathematics Series* 27-2004.

[38] Maccheroni, F., Marinacci, M., Rustichini, A., (2006). Dynamic variational preferences. *Journal of Economic Theory* 128, 4-44.

[39] Pardoux, E., Peng, S., (1990). Adapted solution of a backward stochastic differential equation. *System and Control Letters* 14, 1, 55-61.

[40] Peng, S., (1997). Backward SDE and related g-expectations. In: N. El Karoui and L. Mazliak (eds.), *Backward Stochastic Differential Equations, Pitman Research Notes Mathematics Series* 364, Longman, Harlow, 141-159.

[41] Peng, S., (2004). *Non-linear expectations, non-linear evaluations and risk measures.* Lecture Notes in Mathematics 1856. Springer.

[42] Riedel, F., (2004). Dynamic coherent risk measures. *Stochastic Processes and their Application* 112, 185-200.

[43] Rockafellar, R.T., Uryasev, A., Zabarankin, M., (2006). Generalized deviations in risk analysis. *Stochastic and Finance* 10, 51-74.

[44] Roorda, B., Schumacher, J.M., (2007). Time consistency conditions for acceptability measures, with an application to Tail Value at Risk. *Insurance: Mathematics and Economics* 40, 209-230.

[45] Roorda, B., Schumacher, J.M., Engwerda, J., (2005). Coherent acceptability measures in multiperiod models. *Mathematical Finance* 15, 589-612.

[46] Rosazza Gianin, E., (2006). Risk measures via g-expectations. *Insurance: Mathematics and Economics* 39, 19-34.

[47] Ruszczyński, A., Shapiro, A., (2006a). Optimization of convex risk functions. *Mathematics of Operations Research* 31, 433-452.

[48] Ruszczyński, A. Shapiro, A., (2006b). Conditional risk mappings. *Mathematics of Operations Research* 31, 544-561.

[49] Stadje, M., (2009). Dynamic Risk Measures and BSDEs: From Discrete to Continuous Time. *PhD Thesis*, Princeton.

[50] Wang, S., Young, V.R., Panjer, H.H., (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics* 21, 173-183.

[51] Zalinescu, C., (2002). *Convex Analysis in General Vector Spaces*, World Scientific.