Distribution-Invariant Dynamic Risk Measures

Stefan Weber*

Humboldt-Universität zu Berlin

November 24, 2003; this version December 4, 2003

Abstract

The paper provides an axiomatic characterization of dynamic risk measures for multi-period financial positions. For the special case of a terminal cash flow, we require that risk depends on its conditional distribution only. We prove a representation theorem for dynamic risk measures and investigate their relation to static risk measures. Two notions of dynamic consistency are proposed. A key insight of the paper is that dynamic consistency and the notion of "measure convex sets of probability measures" are intimately related. Measure convexity can be interpreted using the concept of compound lotteries. We characterize the class of static risk measures that represent consistent dynamic risk measures. It turns out that these are closely connected to shortfall risk. Under weak additional assumptions, static convex risk measures coincide with shortfall risk, if compound lotteries of acceptable respectively rejected positions are again acceptable respectively rejected. This result implies a characterization of dynamically consistent convex risk measures.

Key words: Dynamic risk measure, capital requirement, measure of risk, dynamic consistency, measure convexity, shortfall risk

JEL Classification: G18, G11, G28

Mathematics Subject Classification (2000): 91B16, 91B28, 91B30

*Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, email sweber@math.hu-berlin.de. I would like to thank Hans Föllmer, Kay Giesecke, Ulrich Horst, Sebastian Knaak and Alexander Schied for helpful discussions. I acknowledge financial support by Deutsche Forschungsgemeinschaft via Graduiertenkolleg 251 'Stochastische Prozesse und Probabilistische Analysis' and SFB 373 'Quantifikation und Simulation ökonomischer Prozesse'.

1 Introduction

The quantification of the risk of financial positions is a key task for both financial institutions and supervising authorities. Risk management and financial regulation relies on the proper assessment of downside risk. Since traditional approaches – such as *value at risk* – do in general not encourage diversification of positions, alternative risk measures need to be designed and investigated. In the context of static financial positions economically meaningful axioms were proposed in the seminal paper by Artzner, Delbaen, Eber & Heath (1999). The original definition has been relaxed in many directions, and various robust representation results for risk measures have been obtained (see e.g. Föllmer & Schied (2002*a*), Föllmer & Schied (2002*b*), Delbaen (2002)). Risk measures for topological vector spaces were considered by Jaschke & Küchler (2001) and Frittelli & Rosazza (2002). For excellent overviews on static risk measures, we refer to Föllmer & Schied (2002*c*), Delbaen (2000) and Scandolo (2003).

While the theory of static risk measures is already well developed, sophisticated risk management and financial regulation requires *dynamic* risk measures for *dynamic* financial positions. Monetary measures of downside risk must evaluate the total risk of both the terminal and all intermediate cash flows. The measurements must consistently be updated, as new information becomes available. In the current paper, we suggest an axiomatically well-founded model for dynamic risk measures of dynamic cash flows in discrete time. As in the static case, the measurement can be interpreted as a capital requirement that must be invested in a risk-free financial instrument until a terminal date. In contrast to most of the literature, we do *not* require that the risk measure is convex in the sense of Föllmer & Schied (2002*c*).

For certain dynamic risk measures we prove a simple representation theorem in terms of static distribution-invariant risk measures. Besides standard conditions known from the static case, the essential axioms are roughly the following:

(1) Agents have access to a market of risk-free bonds. The risk of two positions is equal at the current date, if they can completely be transformed into each other by trading in the bond market in the future.

(2) Whether or not a terminal position has positive risk, depends only on its conditional distribution.

We propose two notions of dynamic consistency for such risk measures, namely acceptance and rejection consistency. We call a dynamic risk measure acceptance consistent (resp. rejection consistent), if it satisfies the following condition: If a position is acceptable (resp. not acceptable) in the future for sure, then it is acceptable (resp. not acceptable) today. It is shown that dynamic consistency is closely related to properties of the acceptance and rejection sets of the representing static risk measures. Here, we use the concept of measure convex sets known from Choquet theory.

We completely characterize the class of static risk measures that corresponds to consistent dynamic risk measures.

Finally, we further investigate these static distribution-invariant risk measures. Both their acceptance and their rejection sets are convex subsets of the space of probability measures. This has a natural interpretation in the context of static financial positions. If two financial positions or lotteries are acceptable (resp. rejected), than any compound lottery that randomizes over the positions is again acceptable (resp. rejected). Under additional topological conditions, we prove that risk measures with such acceptance and rejection sets coincide exactly with the well-known shortfall risk, if they are convex in the sense of Föllmer & Schied (2002*c*). This result can then be applied to dynamically consistent, convex risk measures.

There are many ways to introduce risk measures in a dynamic setting. Most approaches in the literature generalize the static results on coherent or convex risk measures. In contrast, we focus on distribution invariance and the connection between dynamic consistency and measure convexity. This implies the close link between shortfall risk on the one hand, and dynamic consistency, convexity and distribution-invariance on the other hand.

The axiomatic approach of Riedel (2002) is related to the current paper. He analyzes dynamic coherent risk measures for financial positions on a finite probability space. Under a strong dynamic consistency axiom, he obtains a robust representation of coherent, dynamically consistent risk measures. The notions of dynamic consistency in the context of risk measures go back to Wang (1996) and Wang (1999).

Artzner, Delbaen, Eber, Heath & Ku (2003) consider financial processes as random variables on an extended state space including dates in time. This allows them to employ the standard approach for static coherent risk measures and to obtain a robust representation. They establish a connection between time consistency, stability of test probabilities and Bellman's principle, see also Delbaen (2003). The approaches of Riedel (2002) and Artzner et al. (2003) are related to the analysis of multiple priors in decision theory, see e.g. Epstein & Schneider (2003). Convex and coherent risk measures for continuous-time processes are investigated by Cheridito, Delbaen & Kupper (2003). An axiomatic analysis of convex, conditioned risk measures can be found in Detlefsen (2003) and Scandolo (2003).

We impose a special type of distribution invariance on dynamic risk measures. In the static context, coherent and convex distribution-invariant risk measures have been investigated by Kusuoka (2001), Carlier & Dana (2003), and Kunze (2003). These can be represented in terms of robust mixtures of average value at risk or upper envelops of Choquet integrals with respect to distortions of probability measures.

The paper is organized as follows. In Section 2 we propose an axiomatic characterization of dynamic risk measures. In Section 3, we investigate static risk measures considered as functionals on the space of probability measures, and prove a simple representation theorem for dynamic risk measures in terms of static risk measures. Dynamic consistency conditions and locally measure convex sets of probability measures are considered in Section 4. In Section 5 we investigate the close link of dynamic consistency and shortfall risk. Section 6 concludes. All proofs are given in the appendix.

2 An Axiomatic Characterization of Risk Dynamics

We consider time periods $t = 0, 1, \dots, T$. The state space $(\Omega, \mathcal{F}, P)$ is a standard Borel probability space. $(\mathcal{F}_t)_{t=0,1,\dots,T}$ denotes a filtration, modelling the flow of information. We assume that at time 0 information is trivial, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and that at time T all information is revealed, i.e. $\mathcal{F}_T = \mathcal{F}$.

We intend to construct an axiomatically well-founded model for dynamic risk of financial positions. A dynamic monetary measure of risk is a sequence of mappings $\rho = (\rho_t)_{t=0,1,\dots,T-1}$ evaluating the risk of dynamic cash flows or financial positions $D = (D_t)_{t=0,1,\dots,T}$. The quantity $\rho_t(D)$ is interpreted as a measure of the risk of position D at time t. We suppose that the space of financial positions equals

$$\mathcal{D} = \{(D_t)_{t=0,1,\dots,T} : \; D_t \in L^\infty(\Omega, \mathcal{F}_t, P)\}.$$

The financial position that pays 1 at time t for sure and 0 else will be denoted by

$$e_t = (0,\ 0,\ \dots,\ 0, \underbrace{1}_{t}, \ 0, \dots,\ 0).$$

We assume that agents have access to a market of zero coupon bonds with maturity T. The price of a bond at time t is given by an $\mathcal{F}_t$-measurable random variable P_t^T. Here $P_T^T \equiv 1$, that is, the bond is default free. Considering only a finite time horizon T, we suppose that bond prices are both bounded from below and above, i.e. $P_s^T \in [\epsilon, c]$ for some $0 < \epsilon < c < \infty$. We abstract from trading costs.

2.1 The Axioms

We will assume that a *dynamic risk measure* satisfies the following axioms.

A Adaptedness, Monotonicity and Invariance

(1) *Adaptedness and Boundedness*:
$\rho_t(D) \in L^\infty(\Omega, \mathcal{F}_t, P)$

(2) *Inverse Monotonicity*:
If $D \geq D'$, then $\rho_t(D) \leq \rho_t(D')$.

(3) *Translation-invariance*:
If $Z \in L^\infty(\Omega, \mathcal{F}_t, P)$, then

$$\rho_t\left(D + \frac{Z}{P_t^T} \cdot e_T\right) = \rho_t(D) - Z.$$

A1 ensures that the risk $\rho_t(D)$ of a position D evaluated at time t depends only on information available at time t (adaptedness). Since the position D is bounded, it is reasonable that its risk is also bounded. A2 states that the downside risk of a position decreases, if the payoff of the position increases in all possible scenarios $\omega \in \Omega$.

The axiom of translation-invariance, A3, formalizes the idea that $\rho_t(D)$ is a capital requirement. If an investor invests an amount of Z at time t in a risk-free way until maturity T, her risk is reduced exactly by Z. In particular, A3 implies that

$$\rho_t\left(D + \frac{\rho_t(D)}{P_t^T} \cdot e_T\right) = 0.$$

We will interpret $\rho_t(D)$ as the monetary amount that should be added to D at time t and invested in risk-free bonds until the final date. This makes the position acceptable from the point of view of an investor or regulator, given the information at time t. A position D is acceptable at time t, if its risk $\rho_t(D) \leq 0$. In this case, no positive monetary amount has to be added to the position.

B Independence of the past

If $D_s = D'_s$ for all $s > t$, then $\rho_t(D) = \rho_t(D')$.

B captures the idea that *'sunk costs are sunk.'* When assessing the risk of a position $D \in \mathcal{D}$ at time t, only the future payoffs are taken into account.

C Invariance under adapted transforms

Let $t < u \leq T$, and assume that $Z \in L^\infty(\Omega, \mathcal{F}, P)$ is $\mathcal{F}_u$-measurable. Then

$$\rho_t(D + Z \cdot P_u^T \cdot e_u - Z \cdot e_T) = \rho_t(D).$$

Axiom C can be interpreted as follows. An agent holding a financial position D can form a contingent plan to transform D into $D' = D + Z \cdot P_u^T \cdot e_u - Z \cdot e_T$ without facing any risk at time u:

- Sell Z zero-coupon bonds at time u.
- Pay Z to the bond owners at time T.

Vice versa, an agent holding D' can form a contingent plan to transform D' into D without facing any risk at time u by following the reversed strategy. For the agent the realization of these contingent plans is clearly feasible at the current date t, but it is also still feasible at the later date $t+1$, since u is *strictly* bigger than t. Hence, both positions D and D' are equivalent for the agent at least until date $t+1$.

Thus, *before* time $t+1$ they should have the same risk.[1] In particular, the relation $\rho_t(D) = \rho_t(D')$ should hold.

From the viewpoint of a regulator the same reasoning applies. It is not necessary to impose different monetary requirements on the positions D and D' already at time t, if they can be transformed into each other at a later date without incurring any cost.[2]

Definition 2.1. *A mapping* $\rho = (\rho_t)_{t=0,1,\dots,T-1} : \mathcal{D} \times \Omega \to \mathbb{R}^T$ *is a* dynamic risk measure *if it satisfies the axioms A1, A2, A3, B and C.*

2.2 Distribution-Invariance

Let ρ be a dynamic risk measure. We define the *acceptance indicator* $a = (a_t)_{t=0,1,\dots,T-1}$ of ρ by

$$a_t(D)(\omega) := \mathbf{1}_{(-\infty,0]}(\rho_t(D)(\omega)).$$

If $a_t(D) = 1$, at date t the risk of D is less or equal to 0 and no positive monetary amount has to be added to D to make it acceptable. Conversely, if $a_t(D) = 0$, a positive monetary amount must be added to D to make the position acceptable at date t.

We denote by $\mathcal{M}_{1,c}(\mathbb{R})$ the space of *probability measures on the real line* with compact support. If Y is a real-valued random variable defined on $(\Omega, \mathcal{F}, P)$, we denote by $\mathcal{L}(Y|\mathcal{F}_t)$ the *regular conditional distribution* of Y given $\mathcal{F}_t$.[3]

Definition 2.2. *The dynamic risk measure* ρ *is called* distribution-invariant at maturity *or* M-invariant *if there exists a measurable mapping*

$$H_t : \mathcal{M}_{1,c}(\mathbb{R}) \to \{0,1\}$$

such that for all terminal positions $D = D_T \cdot e_T \in \mathcal{D}$

$$a_t(D) = H_t(\mathcal{L}(D_T|\mathcal{F}_t)).$$

[1] In Axiom C we state that risk is invariant for positions that can be transformed into each other using zero-coupon bonds. One could argue that risk should also be invariant under a more general class of transformations involving possibly other financial instruments. Observe that such an approach would *add* more restrictions on the risk measure, thus *decrease* the level of generality of the analysis.

[2] The intuition behind *invariance under adapted transforms* can be illustrated by the following example. On Monday a lady buys in a supermarket a bottle of red wine for a party on Saturday - not knowing whether it is sweet or dry. The bottle is labelled D. The supermarket sells also red wine with label D' of opposite type. A day later she gets the information which wine is dry and which is sweet, and she may exchange the bottle against a bottle of type D' if she likes to do so. If it is not costly to go to the supermarket and to buy or exchange goods, *on Monday* the evaluation of bottles with label D or D' should be the same.

[3] Properties of regular conditional distributions are stated in the appendix.

M-invariance formalizes the following idea. The purpose of a risk measure is to quantify the downside risk of a financial position. If a financial institution evaluates the risk of a fixed financial cash flow Z to be paid at a fixed reference point in time T, it is reasonable to assume that acceptability should depend only on the conditional distribution of Z given the present information. The use of conditional distributions formalizes the idea that information is processed in a Bayesian fashion.

Of course, if we do *not* fix Z assuming instead that Z is invested into some financial asset or that Z is a position in a larger portfolio, *then* total risk is determined by the conditional distributions and the dependence structure of *all* financial random variables involved. But, if we would like to evaluate a *fixed Z alone*, downside risk should be understood as a property of its conditional distribution only.

3 Representation of Distribution-Invariant Risk

Dynamic M-invariant risk measures can be represented in terms of static distribution-invariant risk measures. This fact is indeed not surprising, and we will state the exact result in Theorem 3.4. The result is useful for the construction of examples of dynamic risk measures. Moreover, dynamic consistency which will be investigated in Section 4 can be characterized via properties of the representing static risk measures.

3.1 Static Distribution-Invariant Risk Measures

Most of the literature on static and dynamic risk measures focuses on coherence and convexity. In such a context it is useful to define risk measures as functionals on a space of financial positions. In contrast, in the current paper issues like distribution-invariance and dynamic consistency are crucial, and it will be convenient to interpret static distribution-invariant risk measures as functionals on probability measures. On the space $\mathcal{M}_{1,c}(\mathbb{R})$ of probability measures on the real line with compact support a partial order $\leq$ is given by *stochastic dominance.*

Definition 3.1. *A mapping $\Theta : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is called a* risk measure *if it satisfies the following conditions for all $\mu, \nu \in \mathcal{M}_{1,c}(\mathbb{R})$:*

- Inverse Monotonicity: *If $\mu \leq \nu$, then $\Theta(\mu) \geq \Theta(\nu)$.*
- Translation Invariance: *If $m \in \mathbb{R}$, then $\Theta(T_m\, \mu) = \Theta(\mu) - m$.*
  Here, for $m \in \mathbb{R}$ the translation operator T_m is given by $(T_m\mu)(\cdot) = \mu(\cdot - m)$.

Inverse monotonicity captures the intuition that risk decreases, if a financial position is concentrated on larger values. Translation invariance formalizes the idea that the risk of a position is actually a monetary requirement: if a monetary amount m is added to the position μ, its risk is decreased by the same amount.

We introduced static risk measures as functionals on the space of probability measures on the real line, while the classical literature on risk measures investigates functionals on spaces of financial positions. The two notions are equivalent in the following sense:

Suppose that $(\Omega', \mathcal{F}', P')$ is an atomless probability space, and let $L^\infty(\Omega', \mathcal{F}', P')$ be a space of financial positions. If $\Theta : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is a risk measure in the sense of Definition 3.1, then $\Theta'(X) = \Theta(\mathcal{L}(X))$ defines a distribution-invariant risk measure on $L^\infty(\Omega', \mathcal{F}', P')$. Conversely, if Θ' is a distribution-invariant risk measure on $L^\infty(\Omega', \mathcal{F}', P')$, then $\Theta(\mu) = \Theta'(X)$ for some $X \sim \mu$ defines a risk measure on $\mathcal{M}_{1,c}(\mathbb{R})$ in the sense of Definition 3.1.

This identification helps to derive properties of risk measures on $\mathcal{M}_{1,c}(\mathbb{R})$ from the classical case. In Appendix A.2 we derive that any risk measure on $\mathcal{M}_{1,c}(\mathbb{R})$ is Lipschitz-continuous with respect to a particular Vasserstein metric. This implies, in particular, that risk measures are measurable functionals with respect to the Borel-σ-algebra of the weak topology.

Acceptance sets on the level of probability distributions can be defined by

$$\mathcal{N}_\Theta = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \ \Theta(\mu) \leq 0\}.$$

For any given risk measure, the acceptance set consists of the probability distributions with non positive risk. Conversely, as in the case of financial positions, acceptance sets may be used to define corresponding risk measures. The following lemma is a simple corollary of the well-known results on classical risk measures, see e.g. Propositions 4.5 and 4.6 in Föllmer & Schied (2002*c*).

Lemma 3.2. *Assume that $\mathcal{N} \subseteq \mathcal{M}_{1,c}(\mathbb{R})$ is non-empty, and satisfies the following two conditions:*

$$\inf\{m \in \mathbb{R} : \quad \delta_m \in \mathcal{N}\} > -\infty. \tag{1}$$

$$\mu \in \mathcal{N}, \ \nu \in \mathcal{M}_{1,c}(\mathbb{R}), \ \nu \geq \mu \ \Rightarrow \ \nu \in \mathcal{N}. \tag{2}$$

Then $\mathcal{N}$ induces a risk measure Θ by

$$\Theta(\mu) = \inf\{m \in \mathbb{R} : \quad T_m(\mu) \in \mathcal{N}\}.$$

$\mathcal{N}$ is included in the acceptance set of Θ.

Recall that the measure of risk Θ' on the space $L^\infty(\Omega', \mathcal{F}', P')$ is called *convex*, if $\Theta'(\alpha X + (1-\alpha)Y) \leq \alpha\Theta'(X) + (1-\alpha)\Theta'(Y)$ for all $X, Y \in L^\infty(\Omega', \mathcal{F}', P')$, $\alpha \in [0,1]$. Θ' is called *positively homogenous*, if $\Theta'(\lambda X) = \lambda\Theta'(X)$ for all $X \in L^\infty(\Omega', \mathcal{F}', P')$ and $\lambda \geq 0$. The risk measure is *coherent*, if it is both convex and positively homogenous. In the next definition[4] we introduce the notions of convexity and coherence for risk measures on $\mathcal{M}_{1,c}(\mathbb{R})$ employing the correspondence to the classical case.

[4]In the Appendix we will show that the concepts of *convexity* and *coherence* of risk measures on $\mathcal{M}_{1,c}(\mathbb{R})$ are indeed well-defined.

Definition 3.3. *Let Θ and Θ' be risk measures as defined above. We say that Θ is* convex *(resp.* coherent*) if Θ' is convex (resp. coherent).*

Under additional continuity conditions, static distribution-invariant risk measures can be represented as robust mixtures of average value at risk and as upper envelops of Choquet integrals with respect to distortions of probability measures. Such characterizations of convex and coherent risk measures follow from results of Kusuoka (2001), Carlier & Dana (2003), and Kunze (2003).

3.2 A Simple Representation Theorem

The following representation characterizes M-invariant dynamic risk measures in a simple way. All proofs are given in the appendix. To keep the notation simple, we denote by

$$\mathcal{T}_t(D) := \mathcal{L}\left(\sum_{u=t+1}^{T} \frac{D_u}{P_u^T}\middle|\mathcal{F}_t\right)$$

the conditional distribution of a specific terminal position associated with $D \in \mathcal{D}$.

Theorem 3.4. *Assume that the probability space is rich in the sense that there exists a unif(0,1)-distributed random variable*[5] *independent of $\mathcal{F}_{T-1}$. Then an M-invariant dynamic risk measure ρ can be represented by*

$$\rho_t(D) = P_t^T \cdot \Theta_t\left[\mathcal{T}_t(D)\right]. \tag{3}$$

Here, Θ_t is a static risk measure considered as a functional on probability measures on $\mathbb{R}$. The risk measures Θ_t in the representation are unique, and the acceptance set of Θ_t is given by

$$\mathcal{N}_t = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \; H_t(\mu) = 1\}. \tag{4}$$

If the dynamic risk measure ρ is *positively homogeneous*, i.e. $\rho_t(\alpha \cdot D) = \alpha \cdot \rho_t(D)$ for $\alpha \in L^\infty(\Omega, \mathcal{F}_t, P)$, then the representing measures Θ_t are *positively homogeneous* and the representation becomes:

$$\rho_t(D) = \Theta_t\left[\mathcal{L}\left(\sum_{u=t+1}^{T} \frac{P_t^T}{P_u^T} \cdot D_u\middle|\mathcal{F}_t\right)\right].$$

If interest rates are deterministic, this representation of *positively homogeneous* risk measures involves only discounted positions. This parallels the results of Riedel (2002) on *coherent* dynamic risk measures on finite probability spaces.

The next lemma states the converse of Theorem 3.4: if the components of ρ are defined as in (3), then ρ is an M-invariant dynamic risk measure.

[5] In Theorem 3.4, Corollary 4.2, Theorem 4.4, and Theorem 4.5 we assume that the underlying probability spaces are rich in an appropriate sense. We formulate these requirements in terms of unif$(0,1)$-distributed random variables. This special assumption on the distribution is not necessary and can be relaxed. Instead, it is equivalent to assume the existence of an arbitrary *continuous* distribution.

Lemma 3.5. *Let $(\Theta_t)_{t=0,1,\ldots,T-1}$ be a sequence of static risk measures as introduced in Definition 3.1. Then (3) defines an M-invariant dynamic risk measure.*

Remark 3.6. *At a given time t the positions $D \in \mathcal{D}$ and $\sum_{u=t+1}^{T} \frac{D_u}{P_u^T} \cdot e_T$ have the same risk. This is implied by axioms B and C, namely by invariance under adapted transforms and independence of the past. The risk of D is then calculated by discounting the static risk of the conditional distribution of the terminal payment $\sum_{u=t+1}^{T} \frac{D_u}{P_u^T}$.*

This result can be generalized in the following way. Instead of requiring axioms B and C, we could assume that at a given time t the position $D \in \mathcal{D}$ has the same risk as a terminal position $T_t(D) \cdot e_T$, where $T_t(D) \in \mathcal{L}^\infty(\Omega, \mathcal{F}, P)$. Here, we suppose that on terminal positions the mapping $D \mapsto T_t(D) \cdot e_T$ is the identity. Define $\mathcal{T}_t(D) := \mathcal{L}(T_t(D)|\mathcal{F}_t)$. Then Theorem 3.4 is still true. If additionally the mappings T_t are monotone increasing on $\mathcal{D}$, the same applies to Lemma 3.5 and the results of Sections 4 and 5. This generalization is important, if due to liquidity risk it is more expensive to transfer large negative amounts to the terminal date than small negative amounts.

4 Dynamic Consistency

The axioms A, B, and C describe the properties of the components ρ_t of the risk measure ρ, but do not require any consistency of risk evaluated at different dates. This fact is also apparent from Theorem 3.4 and Lemma 3.5: the representing static risk measures Θ_t can arbitrarily be chosen for different values of t. In this section we will investigate the implications of consistency requirements in time.

4.1 Representation of Consistent Risk Measures

Definition 4.1. *A dynamic risk measure ρ is*

- acceptance consistent, *if*

$$a_{t+1}(D) \equiv 1 \quad \Rightarrow \quad a_t(D - D_{t+1} \cdot e_{t+1}) \equiv 1,$$

- rejection consistent, *if*

$$a_{t+1}(D) \equiv 0 \quad \Rightarrow \quad a_t(D - D_{t+1} \cdot e_{t+1}) \equiv 0.$$

Here, equality is always understood P-almost surely.

Acceptance consistency captures the following intuition. If a position D is acceptable at the date $t+1$ irrespectively of actual scenario $\omega \in \Omega$, then D should also be accepted at the earlier time t if we neglect the payment at date $t+1$. This payment is not taken into consideration in the definition of consistency, because it does never enter the risk evaluation at time $t+1$ by the axiom of independence of the past. In an analogous manner, *rejection consistency* states the idea that a position should already be rejected at time t if we neglect the payment at $t+1$ and the position is rejected at the later date $t+1$ in any scenario $\omega \in \Omega$.

The consistency conditions have implications for the representation of a distribution-invariant dynamic risk measure given by

$$\rho_t(D) = P_t^T \cdot \Theta_t \left[\mathcal{T}_t(D)\right].$$

Let $\mathcal{N}_t \subseteq \mathcal{M}_{1,c}(\mathbb{R})$ be the acceptance set of the static risk measure Θ_t. Assume that the probability space is rich in the sense that there exists a unif(0,1)-distributed random variable independent of $\mathcal{F}_{T-1}$. Then the following holds:[6]

- If ρ is acceptance consistent, then $\mathcal{N}_{t+1} \subseteq \mathcal{N}_t$.
- If ρ is rejection consistent, then $\mathcal{N}_{t+1} \supseteq \mathcal{N}_t$.

If both consistency conditions are satisfied, we obtain the following corollary.

Corollary 4.2. *Assume that the probability space is rich in the sense that there exists a unif(0,1)-distributed random variable independent of $\mathcal{F}_{T-1}$. Let the M-invariant dynamic risk measure ρ be both acceptance and rejection consistent. Then ρ can be represented by*

$$\rho_t(D) = P_t^T \cdot \Theta \left[\mathcal{T}_t(D)\right]$$

Here, Θ is a unique static risk measure considered as a functional on probability measures on $\mathbb{R}$ with acceptance set

$$\mathcal{N} = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \; H_t(\mu) = 1\} \qquad (t = 0, 1, \ldots, T-1).$$

4.2 Consistency and mixtures of distributions

According to Corollary 4.2 a dynamic risk measure can be represented by one universal static risk measure, if it is both acceptance and rejection consistent. In the following theorem we take the opposite point of view asking the question:
If a dynamic risk measure can be represented by a single static risk measure - what are the properties of the static risk measure, in case the dynamic risk measures satisfies consistency properties?

It turns out that this question can be answered employing the notion of mixtures of probability measures. The following definition introduces the appropriate concept, cf. Winkler (1985).

Definition 4.3. *Let $\mathcal{C}$ be a measurable subset of $\mathcal{M}_{1,c}(\mathbb{R})$. We say that $\mathcal{C}$ is* locally measure convex *if for all $c \in \mathbb{R}$ and any probability measure γ on $\mathcal{C} \cap \mathcal{M}_1([-c, c])$ the mixture $\int \nu \gamma(d\nu)$ is again an element of $\mathcal{C}$.*

[6] The proof is given in Section A.6 (Proof of Corollary 4.2).

The last definition simply formalizes the notion of measure convex sets of probabilities in the context of measures with bounded support. The next theorem gives a first answer to our question.

Theorem 4.4. *Let Θ be a static risk measure, and let $\mathcal{N} \subseteq \mathcal{M}_{1,c}(\mathbb{R})$ be its acceptance set. Then*

$$\rho_t(D) = P_t^T \cdot \Theta\left[\mathcal{T}_t(D)\right]$$

defines an M-invariant dynamic risk measure. If $\mathcal{N}$ is locally measure convex, then ρ is acceptance consistent. If $\mathcal{N}^c$ is locally measure convex, then ρ is rejection consistent.

The characterization of consistency in terms of the acceptance sets of the representing risk measure and mixtures of probability measures can be strengthened if the underlying probability space is rich enough.

Theorem 4.5. *Assume that the probability space is rich in the sense that there exist both a unif(0,1)-distributed random variable independent of $\mathcal{F}_{T-1}$, and a unif(0,1)-distributed, $\mathcal{F}_{T-1}-$measurable random variable independent of $\mathcal{F}_{T-2}$. Assume again that the dynamic risk measure ρ is represented as in Theorem 4.4.*

Then ρ is acceptance consistent, if and only if $\mathcal{N}$ is locally measure convex. Analogously, ρ is rejection consistent, if and only if $\mathcal{N}^c$ is locally measure convex.

4.3 Examples

Theorem 4.4 and Theorem 4.5 are very useful when constructing consistent dynamic risk measures. Examples for static risk measures which induce an acceptance and rejection consistent dynamic risk measure include the negative expected value, the worst-case measure, value at risk, and shortfall risk.

Example 4.6 (Negative expected value, Worst-case measure).
The negative expected value *is given by*

$$\Theta(\mu) = -\int_{\mathbb{R}} x\mu(dx).$$

The worst-case measure *is defined as*

$$\Theta(\mu) = -\inf\{y \in \mathbb{R} : \ \mu(-\infty, y) > 0\}.$$

In both case, the following holds: First, the acceptance set $\mathcal{N} = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \Theta(\mu) \leq 0\}$ and the rejection set $\mathcal{N}^c$ are locally measure convex. Hence, Θ induces an acceptance and rejection consistent dynamic risk measure ρ. Second, Θ is a coherent risk measure. Thus, the components of the dynamic risk measure ρ are coherent on $\mathcal{D}$, that is for $t = 0, 1, \ldots, T-1$ the components satisfy both convexity and positive homogeneity:

- *Convexity:*

$$\rho_t(\alpha D + (1-\alpha)G) \;\leq\; \alpha\rho_t(D) \;+\; (1-\alpha)\rho_t(G)$$
$$(\alpha \in L^\infty(\Omega, \mathcal{F}_t, P),\; 0 < \alpha < 1,\;\; D, G \in \mathcal{D}).$$

- *Positive homogeneity:*

$$\rho_t(\lambda \cdot D) \;=\; \lambda \cdot \rho_t(D) \qquad\qquad (\lambda \in L^\infty(\Omega, \mathcal{F}_t, P),\; \lambda \geq 0,\;\; D \in \mathcal{D}).$$

Example 4.7 (Value at risk).
Value at risk *at level* $\alpha \in [0, 1)$ *is defined as*

$$\begin{aligned}\Theta(\mu) &= -\inf\{y \in \mathbb{R} :\; \mu(-\infty, y] > \alpha\} \\ &= -\sup\{y \in \mathbb{R} :\; \mu(-\infty, y) \leq \alpha\} \\ &= \;\inf\{y \in \mathbb{R} :\; \mu(-\infty, -y) \leq \alpha\}.\end{aligned}$$

The acceptance set $\mathcal{N} = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) :\; \mu(-\infty, 0) \leq \alpha\}$ *and the rejection set* $\mathcal{N}^c$ *are locally measure convex. Hence,* Θ *induces an acceptance and rejection consistent dynamic risk measure* ρ. Θ *is not a convex risk measure. Thus, the time components of the dynamic risk measure* ρ *are not convex on* $\mathcal{D}$.

Example 4.8 (Shortfall risk).
Let $\ell : \mathbb{R} \to \mathbb{R}$ *be a convex loss function, i.e. an increasing, non constant and convex function. Assume that* z *is an interior point of the range of* ℓ.
We define an acceptance set

$$\mathcal{N} = \left\{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) :\; \int \ell(-x)\mu(dx) \leq z\right\}.$$

$\mathcal{N}$ *induces the short-fall risk measure* Θ *by*

$$\Theta(\mu) = \inf\{m \in \mathbb{R} :\; T_m\mu \in \mathcal{N}\}.$$

Here, for $m \in \mathbb{R}$ *the translation operator* T_m *is given by*

$$(T_m\mu)(\cdot) = \mu(\cdot - m).$$

The induced dynamic risk measure will be denoted by ρ.
Shortfall risk has the following properties:

(1) Acceptance and rejection set are locally measure convex. Hence, ρ *is acceptance and rejection consistent.*

(2) Θ *is convex. Thus, the components of* ρ *are convex on* $\mathcal{D}$.

An exponential loss function

$$\ell(x) = \exp(ax) \qquad (a > 0)$$

leads to the special case of the entropic risk measure

$$\Theta(\mu) = \frac{1}{a}\left(\log \int \exp(-ax)\mu(dx) - \log z\right).$$

5 Consistency, Compound Lotteries, and Shortfall Risk

The static risk measures representing dynamically consistent risk measures are closely related to shortfall risk. Theorem 5.3 will demonstrate the close link which relies on a weak closure property of the acceptance set. Before stating the theorem we need to introduce topologies on $\mathcal{M}_{1,c}(\mathbb{R})$ that allow us to deal with integrals against unbounded test functions.
For a fixed continuous function

$$\psi : \mathbb{R} \to [1, \infty)$$

we denote by C^ψ the vector space of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ for which we can find a constant $c \in \mathbb{R}$ such that for all $x \in \mathbb{R}$,

$$|f(x)| \leq c \cdot \psi(x).$$

ψ is called a *gauge function.* $\mathcal{M}_c^+(\mathbb{R})$ designates the space of finite measures with compact support.

Definition 5.1. *The ψ-weak topology on the set $\mathcal{M}_c^+(\mathbb{R})$ is the initial topology of the family $\mu \mapsto \int f(x)\mu(dx)$ $(\mu \in \mathcal{M}_c(\mathbb{R}),\ f \in C^\psi)$.*

In other words, the ψ-weak topology is the weakest topology on $\mathcal{M}_c^+(\mathbb{R})$ for which all mappings $\mu \mapsto \int f(x)\mu(dx)$ $(\mu \in \mathcal{M}_c(\mathbb{R}))$ with $f \in C^\psi$ are continuous. It is *finer* than the weak topology. Convergence of sequences of measures can be characterized as follows:

Lemma 5.2. *A sequence of measures $(\mu_n)_{n\in\mathbb{N}}$ in $\mathcal{M}_c^+(\mathbb{R})$ converges ψ-weakly to $\mu \in \mathcal{M}_c^+(\mathbb{R})$ if and only if*

$$\int f d\mu_n \longrightarrow \int f d\mu$$

for every measurable function f which is μ-almost everywhere continuous and for which exists a constant $c \in \mathbb{R}$ such that $|f| \leq c \cdot \psi$ μ-almost everywhere.

5.1 Static Risk Measures

After these preparations we are now able to state the theorem which links shortfall risk and static risk measures representing consistent dynamic risk measures. Recall that a loss function is a non decreasing function which is not identically constant.

Theorem 5.3. *Let Θ be a risk measure on $\mathcal{M}_{1,c}(\mathbb{R})$. Assume that there exists $x \in \mathbb{R}$ with $\delta_x \in \mathcal{N}$ such that for $y \in \mathbb{R}$, $\delta_y \in \mathcal{N}^c$,*

$$(1 - \alpha)\delta_x + \alpha\delta_y \in \mathcal{N} \tag{5}$$

for sufficiently small $\alpha > 0$. Then the following statements are equivalent:

(1) Both the acceptance set $\mathcal{N}$ and the rejection set $\mathcal{N}^c$ of Θ are convex, and $\mathcal{N}$ is ψ-weakly closed for some gauge function $\psi : \mathbb{R} \to [1, \infty)$.

(2) There exists a left-continuous loss function $l : \mathbb{R} \to \mathbb{R}$ and a scalar $z \in \mathbb{R}$ in the interior of the convex hull of the range of l such that

$$\mathcal{N} = \left\{ \mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \int l(-x)\mu(dx) \leq z \right\}.$$

The convexity of the acceptance and rejection sets has a natural interpretation in the context of static financial positions. If two probability measures μ and ν are acceptable (resp. rejected), than for $\alpha \in [0, 1]$ the compound lottery $\alpha\mu + (1 - \alpha)\nu$, that randomizes over μ and ν, is also acceptable (resp. rejected).

Remark 5.4.
The risk measures characterized in the last theorem are closely connected to classical utility theory of von Neumann and Morgenstern. Setting $u(x) := -l(-x)$, we can interpret u as a Bernoulli utility function. Then, a financial position $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$ is considered acceptable, if its expected utility is larger than z,

$$U(\mu) := \int u(x)\mu(dx) \geq z.$$

Remark 5.5. *The functional $\mu \mapsto \int l(-x)\mu(dx)$ is ψ-weakly continuous for some gauge function ψ, if and only if l is continuous. This follows from the representation of the dual space of $\mathcal{M}_{1,c}(\mathbb{R})$ endowed with the ψ-weak topology, cf. Lemma A.7. Let $\psi \in C(\mathbb{R})$, $\psi \geq |g| + 1$ with $g(x) = l(-x)$ $(x \in \mathbb{R})$. In general, the functional is only lower semicontinuous for the ψ-weak topology.*[7]

Example 5.6. *Condition (5) excludes that Θ equals the worst case measure plus some constant (say r), i.e.*

$$\Theta(\mu) \;=\; r \;-\; \textit{ess inf}\, \mu \qquad\qquad (\mu \in \mathcal{M}_{1,c}(\mathbb{R})).$$

Example 5.7. *For the* negative expected value *the loss function is given by $l(x) = x$ with threshold $z = 0$. For* value at risk *at level $\lambda \in (0, 1)$ the loss function equals $l(x) = \mathbf{1}_{(0,\infty)}$ with threshold $z = \lambda$.* Shortfall risk *is already defined in terms of a loss function; characterizations and specific examples will be discussed below.*

Example 5.8. *For a given level $x \in [0, 1)$, let VaR_x be value at risk at level x as defined in Example 4.7. For $\lambda \in (0, 1)$* average value at risk *at level λ is defined by*

$$AVaR_\lambda(\mu) = \frac{1}{\lambda} \int_0^\lambda VaR_x(\mu)dx, \quad \mu \in \mathcal{M}_{1,c}(\mathbb{R}).$$

The acceptance set of $AVaR_\lambda$ $(\lambda \in (0, 1))$ is not convex as subsets of the space of probability measures. A counterexample is given in the appendix. Hence, $AVaR_\lambda$ does not *satisfy condition (1) of Theorem 5.3, and its acceptance set cannot be represented in terms of a loss function.*

[7] See the proof of Theorem 5.3.

The following corollary connects the preceding results with the classical theory of convex risk measures, cf. Chapter 4.6. in Föllmer & Schied (2002*c*).

Corollary 5.9. *Let Θ be a risk measure on $\mathcal{M}_{1,c}(\mathbb{R})$, and assume that its acceptance set $\mathcal{N}$ is characterized as in condition (2) of Theorem 5.3. Then Θ is convex if and only if the loss function l is convex.*

Theorem 5.3 and Corollary 5.9 imply that any convex risk measure Θ on $\mathcal{M}_{1,c}(\mathbb{R})$ with locally measure convex acceptance and rejection set can be represented as *shortfall risk*, if the acceptance set is ψ-weakly closed for some gauge function. Shortfall risk allows a robust representation in terms of the Fenchel-Legendre transform of the associated loss function.

Lemma 5.10. *Let Θ be shortfall risk as defined in Example 4.8 associated with a convex and continuous loss function l. We denote the Fenchel-Legendre transform of l by*

$$l^*(y) := \sup_{x \in \mathbb{R}} (yx - l(x)).$$

A robust representation of the risk measure is given by

$$\Theta(\mu) = \max_{\nu \in \mathcal{M}_1(\mu)} \left(-\int x\nu(dx) - \alpha(\nu|\mu) \right) \qquad (\mu \in \mathcal{M}_{1,c}(\mathbb{R})).$$

Here, $\mathcal{M}_1(\mu)$ is the set of probability measures which are absolutely continuous with respect to μ. The penalty function α is given by

$$\alpha(\nu|\mu) = \inf_{\lambda > 0} \frac{1}{\lambda} \left(z + \int l^* \left(\lambda \frac{d\nu}{d\mu} \right) d\mu \right) \qquad (\nu \in \mathcal{M}_1(\mu)).$$

Example 5.11. *The special choice of the loss function $l(x) = \exp(\alpha \cdot x)$ is associated with the* entropic risk measure*. In this case, a penalty function can be defined in terms of the relative entropy:*

$$\alpha(\nu|\mu) = \frac{1}{\alpha} \left(H(\nu|\mu) - \log z \right) \qquad (\nu \in \mathcal{M}_1(\mu)).$$

Here, the relative entropy is given by

$$H(\nu|\mu) = \begin{cases} \int \frac{d\nu}{d\mu} \log\left(\frac{d\nu}{d\mu}\right) d\mu & \text{if } \nu \ll \mu, \\ \infty & \text{else.} \end{cases}$$

Example 5.12. *Another example that allows explicit calculations*[8] *is given by the convex loss functional*

$$l(x) = \begin{cases} \frac{1}{p} x^p & \text{if } x \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

[8] See e.g. Föllmer & Schied (2002*c*), Example 4.64.

where $p > 1$. Denoting by $q = p/(p-1)$ the dual coefficient, the Legendre-Fenchel transform is calculated as

$$l^*(y) = \begin{cases} \frac{1}{q}\, y^q & \text{if } y \geq 0, \\ \infty & \text{otherwise.} \end{cases}$$

A penalty function is then given by

$$\alpha^p(\nu|\mu) = (p \cdot z)^{1/p} \left(\int \left(\frac{d\nu}{d\mu} \right)^q d\mu \right)^{1/q} \qquad (\nu \in \mathcal{M}_1(\mu)).$$

The case of classical expected shortfall risk $l(x) = x^+$ is obtained for $p \searrow 1$. A penalty function can be calculated as

$$\alpha(\nu|\mu) \;=\; z \cdot \left\| \frac{d\nu}{d\mu} \right\| \qquad (\nu \in \mathcal{M}_1(\mu)).$$

Finally we consider the case of coherent risk measures.

Corollary 5.13. *Let Θ be a risk measure on $\mathcal{M}_{1,c}(\mathbb{R})$, and assume that its acceptance set $\mathcal{N}$ is characterized as in condition (2) of Theorem 5.3. Then Θ is coherent if and only if $l(x) = z + \alpha x^+ - \beta x^-$ for $\alpha \geq \beta > 0$.*

For coherent measures of risk that satisfy the assumptions of Theorem 5.3 a position is acceptable, if a suitable weighted average of expected gains and expected losses is sufficiently large. In particular, gains and losses can be weighted differently, and the weight of the losses is not smaller than the weight of the gains.

While the conditions given in Theorem 5.3 together with convexity are all highly desirable, the additional requirement of *positive homogeneity* implicit in the notion of *coherence* has frequently been criticized in the literature. It neglects the possibility that risk might grow in a nonlinear fashion, if borrowing constraints and liquidity risk are present.

5.2 Dynamic Risk Measures

The results of the last section can be applied to dynamic risk measures. Dynamic consistency, convexity and a weak closure property imply that a dynamic risk measure can be represented in terms of shortfall risk.

Theorem 5.14. *Assume that the probability space is rich in the sense that there exist both a unif(0,1)-distributed random variable independent of $\mathcal{F}_{T-1}$, and a unif(0,1)-distributed, $\mathcal{F}_{T-1}-$measurable random variable independent of $\mathcal{F}_{T-2}$. Let ρ be an M-invariant dynamic risk measure. We make the following assumptions:*

(1) ρ is acceptance and rejection consistent.

(2) ρ is convex in the sense that for $t = 0, 1, \ldots, T-1$, $\alpha \in (0,1)$, $D, G \in \mathcal{D}$,

$$\rho_t(\alpha D + (1-\alpha)G) \leq \alpha \rho_t(D) + (1-\alpha)\rho_t(G).$$

(3) The set $\mathcal{N} = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \; H_t(\mu) = 1\}$ $(t = 0, 1, \ldots, T-1)$ *is* ψ*-weakly closed for some gauge function* $\psi : \mathbb{R} \to [1, \infty)$.

(4) Assume that there exists $x \in \mathbb{R}$ *with* $\delta_x \in \mathcal{N}$ *such that for* $y \in \mathbb{R}$, $\delta_y \in \mathcal{N}^c$,

$$(1-\alpha)\delta_x + \alpha\delta_y \in \mathcal{N}$$

for sufficiently small $\alpha > 0$.

Then there exists a continuous and convex loss function $l : \mathbb{R} \to \mathbb{R}$ *with associated shortfall risk measure* Θ *on* $\mathcal{M}_{1,c}(\mathbb{R})$ *such that* ρ *can be represented as*

$$\rho_t(D) = P_t^T \cdot \Theta \left[\mathcal{T}_t(D)\right]. \tag{6}$$

From the point of view of an investor or regulator, distribution-invariance at the reference time T, convexity, and dynamic consistency are desirable properties of a dynamic risk measure. The additional requirement on $\mathcal{N}$ to be ψ-weakly closed for some gauge function ψ is very weak and is even economically meaningful: terminal positions which can be approximated by acceptable positions in a rather fine topology are again acceptable. We argue therefore that static shortfall risk provides a good basis for the dynamic evaluation of dynamic financial positions.[9]

6 Conclusion

The paper provides an axiomatic characterization of dynamic risk measures for dynamic cash flows. For the special case of terminal financial positions at a given reference date, we require that the risk measure depends on their conditional distribution only. A key insight of the paper is that dynamic consistency and the notion of measure convex sets of probability measures are intimately related. Measure convexity can be interpreted using the concept of compound lotteries. We characterize the class of static risk measures that represent consistent dynamic risk measures. It turns out that these are closely connected to shortfall risk. Under weak additional assumptions, static convex risk measures coincide with shortfall risk if compound lotteries of acceptable respectively rejected positions are again acceptable respectively rejected. This result implies a characterization of dynamically consistent convex risk measures.

A Appendix

A.1 Regular Conditional Distributions

Regular conditional distributions are a standard tool in probability theory. In this section we recall its definition and results regarding existence and uniqueness.

[9] In case of additional model uncertainty, an investor or regulator should consider robust versions of the results discussed in the current paper. Such an extension is, however, a topic of future research.

Definition A.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let Y be a measurable function on Ω into any measurable space $(T, \mathcal{B})$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then a regular conditional distribution $\mathcal{L}(Y|\mathcal{G})$ of Y given $\mathcal{G}$ is defined as a function from $\Omega \times \mathcal{B}$ into $[0,1]$ such that*

(1) for P-almost all $\omega \in \Omega$, $\mathcal{L}(Y|\mathcal{G})(\omega, \cdot)$ is a probability measure on $\mathcal{B}$.

(2) for each $B \in \mathcal{B}$, $\mathcal{L}(Y|\mathcal{G})(\cdot, B)$ is $\mathcal{G}$-measurable.

(3) for $B \in \mathcal{B}$ and for all $C \in \mathcal{G}$ it holds that

$$\int_C \mathcal{L}(Y|\mathcal{G})(\omega, B) P(d\omega) = \int_C \mathbf{1}_{Y \in B}(\omega) P(d\omega).$$

Theorem A.2. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let Y be a measurable function on Ω into any standard Borel space $(T, \mathcal{B})$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then a regular conditional distribution $\mathcal{L}(Y|\mathcal{G})$ of Y given $\mathcal{G}$ exists. It is unique in the following sense: If $\hat{\mathcal{L}}(Y|\mathcal{G})$ is another regular conditional distribution, then the two laws $\mathcal{L}(Y|\mathcal{G})(\omega, \cdot)$ and $\hat{\mathcal{L}}(Y|\mathcal{G})(\omega, \cdot)$ are equal for P-almost all $\omega \in \Omega$.*

A.2 Vasserstein metric and Lipschitz continuity

Lemma A.3. *Any risk measure $\Theta : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is Lipschitz continuous with respect to the Vasserstein distance V_∞:*

$$|\Theta(\mu) - \Theta(\nu)| \leq V_\infty(\mu, \nu).$$

Here, for $\mu, \nu \in \mathcal{M}_{1,c}(\mathbb{R})$ the Vasserstein distance is defined by

$$V_\infty(\mu, \nu) = \inf \|X - Y\|,$$

where $\|\cdot\|$ denotes the essential supremum and the infimum is taken over all pairs of random variables $X \sim \mu$ and $Y \sim \nu$ on some atomless probability space.

Proof.
Let $\mu, \nu \in \mathcal{M}_{1,c}(\mathbb{R})$ be given. Assume that $X \sim \mu$, $Y \sim \nu$ and $X, Y \in L^\infty(\Omega', \mathcal{F}', P')$ for some probability space $(\Omega', \mathcal{F}', P')$. W.l.o.g we may assume that $(\Omega', \mathcal{F}', P')$ is atomless by identifying every atom with a subinterval of $((0,1), \lambda)$ of appropriate length; this does neither change the joint distribution of (X, Y) nor the norm $\| X - Y \|$. Then by the Lipschitz continuity of Θ' it follows that

$$|\Theta(\mu) - \Theta(\nu)| = |\Theta'(X) - \Theta'(Y)| \leq \|X - Y\|.$$

Note that the Lipschitz continuity of Θ' is a trivial consequence of the monotonicity and translation invariance of Θ', cf. Lemma 4.3 in Föllmer & Schied (2002*c*).
This implies the claim. □

Remark A.4. *For measures on $\mathbb{R}$ the Vasserstein metric V_∞ can be represented in terms of the inverse of the distribution functions (i.e. the quantile functions) of the measures $\mu, \nu \in \mathcal{M}_{1,c}(\mathbb{R})$, cf. Owen (1987). We denote by F_μ^{-1} and F_ν^{-1} the right-continuous inverse of the distribution function of μ and ν, respectively. It holds that*

$$V_\infty(\mu,\nu) = \sup_{0<u<1} |F_\mu^{-1}(u) - F_\nu^{-1}(u)|. \tag{7}$$

For other Vasserstein metrics see Owen (1987) and Rachev (1991).

Lemma A.5. *The V_∞-metric generates the Borel-σ-algebra on $\mathcal{M}_{1,c}(\mathbb{R})$ induced by the weak topology.*

Proof. The quantile function

$$F_\mu^{-1}(u) = q_\mu(u) = \sup\{x : \ \mu(-\infty,x) \le u\} \tag{8}$$

is product measurable on $\mathcal{M}_{1,c}(\mathbb{R}) \times [0,1]$, since the set $\{(\mu,u) : \ \mu(-\infty,x) \le u\}$ is measurable for each x and the supremum in (8) can be restricted to rational x. More precisely, the product measurability is implied by the following identities: for any $z \in \mathbb{R}$ it holds that

$$\begin{aligned} \{(\mu,u) : \ q_\mu(u) \ge z\} \ &= \ \left\{ (\mu,u) : \ \sup_{x\in\mathbb{Q}}\{x : \ \mu(-\infty,x) \le u\} \ge z\} \right\} \\ &= \bigcap_{x\in\mathbb{Q},\, x<z} \{(\mu,u) : \ \mu(-\infty,x) \le u\}. \end{aligned}$$

Now fix $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$. Then the Vasserstein ball $\{\nu : V_\infty(\mu,\nu) < \epsilon\}$ is measurable with respect to the standard σ-algebra, since the supremum in (7) can be restricted to rational u. More precisely, for $u \in [0,1]$ the function $q_\cdot(u) : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is measurable with respect to the standard σ-algebra. Hence,

$$V_\infty(\mu,\cdot) \ = \ \sup_{0<u<1,\, u\in\mathbb{Q}} |q_\mu(u) - q_\cdot(u)|$$

is measurable with respect to the standard σ-algebra. This implies the measurability of the Vasserstein ball.

Hence, the Borel-σ-algebra generated by the V_∞-topology is coarser that the standard σ-algebra. The converse is true, since the Vasserstein topology is finer that the weak topology. □

Corollary A.6. *A risk measure $\Theta : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is measurable with respect to the Borel-σ-algebra on $\mathcal{M}_{1,c}(\mathbb{R})$ generated by the weak topology.*

A.3 Verification of Definition 3.3.

We have to show that the definitions do not depend on the choice of the atomless probability space $(\Omega', \mathcal{F}', P')$. Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ be another atomless probability space, and let Z be a unif$(0,1)$-distributed random variable on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$. Let $X', Y' : \Omega' \to \mathbb{R}$ be two random variables. By Borel's theorem (see e.g. Theorem 2.19 in Kallenberg (1997)) it follows that there exists a measurable mapping $(g_1, g_2) : [0,1] \to \mathbb{R}^2$ such that $(g_1 \circ Z, g_2 \circ Z) \sim (X', Y')$. We set $\hat{X} = g_1 \circ Z$, $\hat{Y} = g_2 \circ Z$.

Now suppose that $\alpha \in (0,1)$, and for random variables $\hat{X}, \hat{Y} : \hat{\Omega} \to \mathbb{R}$,

$$\Theta(\mathcal{L}(\alpha \hat{X} + (1-\alpha)\hat{Y}) \leq \alpha \Theta(\mathcal{L}(\hat{X})) + (1-\alpha)\Theta(\mathcal{L}(\hat{Y})).$$

Let $\alpha \in (0,1)$, and random variables $X', Y' : \Omega' \to \mathbb{R}$ be given. Then there exists random variables $\hat{X}, \hat{Y} : \hat{\Omega} \to \mathbb{R}$ such that $(X', Y') \sim (\hat{X}, \hat{Y})$. We obtain

$$\begin{aligned} \Theta(\mathcal{L}(\alpha X' + (1-\alpha)Y') &= \Theta(\mathcal{L}(\alpha \hat{X} + (1-\alpha)\hat{Y}) \\ \leq \alpha\Theta(\mathcal{L}(\hat{X})) + (1-\alpha)\Theta(\mathcal{L}(\hat{Y})) &= \alpha\Theta(\mathcal{L}(X')) + (1-\alpha)\Theta(\mathcal{L}(Y')). \end{aligned}$$

The same implication holds if we reverse the roles of Ω' and $\hat{\Omega}$. It follows that the definition of convexity of Θ does not rely on the choice of the probability space $(\Omega', \mathcal{F}', P')$. An analogous argument holds for coherence. □

A.4 Proof of Theorem 3.4.

Let $D \in \mathcal{D}$ be given. By independence of the past and invariance under adapted transforms we obtain

$$\rho_t(D) = \rho_t\left(\sum_{u=t+1}^{T} D_u \cdot e_u\right) = \rho_t\left(\sum_{u=t+1}^{T} \frac{D_u}{P_u^T} \cdot e_T\right)$$

Thus, w.l.o.g. we may assume that $D = K \cdot e_T$ with $K \in L^\infty(\Omega, \mathcal{F}, P)$.
For $t = 0, 1, \ldots, T-1$ we define the sets

$$\mathcal{N}_t = \{\mu \in \mathcal{M}_{1,c}(\mathbb{R}) : \ H_t(\mu) = 1\}.$$

We show that $\mathcal{N}_t$ induces a static risk measure.

First, we prove property (1): Let $M' \in L^\infty(\Omega, \mathcal{F}, P)$ be arbitrary. Define

$$M := M' + \frac{\rho_t(M' \cdot e_T) - 1}{P_t^T}.$$

By assumption, P_t^T is bounded away from zero and $\rho_t(M' \cdot e_T) \in L^\infty(\Omega, \mathcal{F}, P)$. Thus, $M \in L^\infty(\Omega, \mathcal{F}, P)$. By translation invariance,

$$\rho_t(M \cdot e_T) = \rho_t\left(M' \cdot e_T + \frac{\rho_T(M' \cdot e_T) - 1}{P_t^T} \cdot e_T\right) = \rho_t(M' \cdot e_T) - \rho_t(M' \cdot e_T) + 1 > 0.$$

Let $m \in \mathbb{R}$, $m \leq -\|M\|_\infty$. By inverse monotonicity, $\rho_t(m \cdot e_T) \geq \rho_t(M \cdot e_T) > 0$. Hence,

$$H_t(\delta_m) = a_t(m \cdot e_T) = 0.$$

This implies that $\inf\{m \in \mathbb{R} : \ \delta_m \in \mathcal{N}_t\} > -\infty$.

Second, we prove property (2): Let $\mu \in \mathcal{N}_t$, $\nu \in \mathcal{M}_{1,c}(\mathbb{R})$, and $\nu \geq \mu$. Since the filtered probability space is rich, there exists a random variable Z uniformly distributed on $(0,1)$ and independent of $\mathcal{F}_{T-1}$. Define $M := q_\mu(Z) \sim \mu$ and $N := q_\nu(Z) \sim \nu$, where q_μ and q_ν are the quantile functions of μ and ν, respectively. Since ν stochastically dominates μ, we have $N \geq M$. By monotonicity, $\rho_t(N \cdot e_T) \leq \rho_t(M \cdot e_T)$. This implies $H_t(\nu) = 1$, since $H_t(\mu) = 1$ by assumption. Hence, $\nu \in \mathcal{N}_t$.

We denote the static risk measure induced by the set $\mathcal{N}_t$ by Θ_t and have to show that

$$\rho_t(D) = P_t^T \cdot \Theta_t(\mathcal{L}(K|\mathcal{F}_t)).$$

By $T : \mathbb{R} \times \mathcal{M}_{1,c}(\mathbb{R}) \to \mathcal{M}_{1,c}(\mathbb{R})$ we denote the translation operator, i.e. $T_r\mu(A) = \mu(A - r)$ for $r \in \mathbb{R}$, $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$ and measurable $A \subseteq \mathbb{R}$.
Since $\rho_t(D) \cdot \left(P_t^T\right)^{-1}$ is $\mathcal{F}_t$-measurable and bounded, we get

$$\frac{\rho_t(D)}{P_t^T} = \text{ess inf } \left\{ m \in L^\infty(\Omega, \mathcal{F}_t, P) : \ \frac{\rho_t(D)}{P_t^T} \leq m \right\}$$

Now let $m \in L^\infty(\Omega, \mathcal{F}_t, P)$ be arbitrary. By translation-invariance,

$$\frac{\rho_t(D)}{P_t^T} - m = \frac{\rho_t(D + m \cdot e_T)}{P_t^T}.$$

Thus,

$$\frac{\rho_t(D)}{P_t^T} \leq m \ \Leftrightarrow \ \rho_t(D + m \cdot e_T) \leq 0 \ \Leftrightarrow \ \mathcal{L}(K + m|\mathcal{F}_t) \in \mathcal{N}_t \ \Leftrightarrow \ T_m\mathcal{L}(K|\mathcal{F}_t) \in \mathcal{N}_t.$$

This implies

$$\frac{\rho_t(D)}{P_t^T} = \text{ess inf } \left\{ m \in L^\infty(\Omega, \mathcal{F}_t, P) : \ \ T_{m(\omega)}\mathcal{L}(K|\mathcal{F}_t)(\omega) \in \mathcal{N}_t \text{ for all } \omega \in \Omega \right\}$$

We have to show that the right hand side equals $\Theta_t(\mathcal{L}(K|\mathcal{F}_t))$:
First, observe that $\Theta_t : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is Lipschitz continuous with respect to the Vasserstein metric V_∞. This implies that $\hat{m} := \Theta_t(\mathcal{L}(K|\mathcal{F}_t)) \in L^\infty(\Omega, \mathcal{F}_t, P)$. Clearly, $T_{\hat{m}(\omega)}\mathcal{L}(K|\mathcal{F}_t)(\omega) \in \mathcal{N}_t$ for all $\omega \in \Omega$. Thus, $\hat{m} \geq \frac{\rho_t(D)}{P_t^T}$.
Second, let $m \in L^\infty(\Omega, \mathcal{F}_t, P)$ such that $T_{m(\omega)}\mathcal{L}(K|\mathcal{F}_t)(\omega) \in \mathcal{N}_t$ for all $\omega \in \Omega$. Since

$$\hat{m}(\omega) = \Theta_t(\mathcal{L}(K|\mathcal{F}_t)(\omega)) = \inf\{r \in \mathbb{R} : \ T_r\mathcal{L}(K|\mathcal{F}_t)(\omega) \in \mathcal{N}_t\},$$

we obtain in particular $\hat{m}(\omega) \leq m(\omega)$ for all $\omega \in \Omega$. Hence $\hat{m} \leq \frac{\rho_t(D)}{P_t^T}$.

Finally, we show that $\mathcal{N}_t$ is indeed the acceptance set of Θ_t and the uniqueness of the representation. Since the probability space is rich, for μ we can find $M \in L^\infty$ with $\mathcal{L}(M|\mathcal{F}_t) = \mu$. Uniqueness is implied by the equality

$$\Theta_t(\mu) = \frac{\rho_t(M \cdot e_T)}{P_t^T}.$$

Moreover, if $\Theta_t(\mu) \leq 0$, then $H_t(\mu) = 1$, thus $\mu \in \mathcal{N}_t$. This implies that $\mathcal{N}_t$ is indeed the acceptance set of Θ_t. □

A.5 Proof of Lemma 3.5.

Adaptedness, inverse monotonicity, and independence of the past are immediate. Boundedness follows from the boundedness assumptions on the bond prices and the Lipschitz continuity of static risk measures with respect to the Vasserstein metric V_∞.

We denote again by $T : \mathbb{R} \times \mathcal{M}_{1,c}(\mathbb{R}) \to \mathcal{M}_{1,c}(\mathbb{R})$ the translation operator. Then translation invariance can be verified as follows. Let $Z \in L^\infty(\Omega, \mathcal{F}_t, P)$. Then

$$\begin{aligned}
\rho_t\left(D + \frac{Z}{P_t^T} \cdot e_T\right) &= P_t^T \cdot \Theta_t\left(\mathcal{L}\left(\sum_{u=t+1}^{T} \frac{D_u}{P_u^T} + \frac{Z}{P_t^T} \middle| \mathcal{F}_t\right)\right) \\
&= P_t^T \cdot \Theta_t\left(T_{\frac{Z}{P_t^T}}\, \mathcal{L}\left(\sum_{u=t+1}^{T} \frac{D_u}{P_u^T} \middle| \mathcal{F}_t\right)\right) = P_t^T \cdot \Theta_t\left(\mathcal{L}\left(\sum_{u=t+1}^{T} \frac{D_u}{P_u^T} \middle| \mathcal{F}_t\right)\right) - Z \\
&= \rho_t(D) - Z
\end{aligned}$$

In order to prove invariance under adapted transforms let $t < v \leq T$, and assume that $Z \in L^\infty(\Omega, \mathcal{F}_v, P)$. Let $D \in \mathcal{D}$ be given, and define $D' = D + Z \cdot P_v^T \cdot e_v - Z \cdot e_T$. The claim follows by observing

$$\sum_{u=t+1}^{T} \frac{D_u}{P_u^T} = \sum_{u=t+1}^{T} \frac{D_u}{P_u^T} + \frac{Z \cdot P_v^T}{P_v^T} - Z = \sum_{u=t+1}^{T} \frac{D'_u}{P_u^T}$$

□

A.6 Proof of Corollary 4.2.

Assume that ρ is acceptance consistent. Let $\mu \in \mathcal{N}_{t+1}$. Since the probability space is rich, there exists a random variable $Z \sim unif(0,1)$ independent of $\mathcal{F}_{T-1}$. We define $K = q_\mu(Z)$ where q_μ is the quantile function of μ. Observe that $\mathcal{L}(K|\mathcal{F}_t) = \mathcal{L}(K|\mathcal{F}_{t+1}) = \mu$. Let $D := K \cdot e_T$. We obtain that

$$1 = H_{t+1}(\mu) = a_{t+1}(D) = a_t(D) = H_t(\mu).$$

Hence, $\mu \in \mathcal{N}_t$.
If ρ is rejection consistent, the proof is analogous.

□

A.7 Proof of Theorem 4.4.

First, ρ defines a M-invariant dynamic risk measure by Lemma 3.5. Second, we prove that ρ is acceptance consistent, if $\mathcal{N}$ is locally measure convex. The case of rejection consistency will then work analogously.
It is not difficult to see that independence of the past and invariance under adapted transforms implies that it suffices w.l.o.g. to investigate terminal positions only, i.e. positions $D \in \mathcal{D}$ of the form $D = K \cdot e_T$ with $K \in L^\infty(\Omega, \mathcal{F}, P)$. By $c \in \mathbb{R}$ we denote some real number such that $K \in [-c, c]$. Define now a kernel K_t from $(\Omega, \mathcal{F}_t)$ to $(\Omega, \mathcal{F})$ such that for measurable $A \subseteq \Omega$,

$$K_t(\omega, A) = P(A|\mathcal{F}_t)(\omega)$$

Set $\mu_s := \mathcal{L}(K|\mathcal{F}_s)$. Then we obtain by disintegration for P-almost every $\omega \in \Omega$ that

$$\mu_t(\omega, \cdot) = \int \mu_{t+1}(\omega', \cdot) K_t(\omega, d\omega')$$

Suppose that $a_{t+1}(D) \equiv 1$. Then $\mu_{t+1}(\omega', \cdot) \in \mathcal{N} \cap \mathcal{M}_1([-c, c])$ for P-almost all $\omega' \in \Omega$. Hence for P-almost all $\omega \in \Omega$,

$$\mu_t(\omega, \cdot) = \int \mu_{t+1}(\omega', \cdot) K_t(\omega, d\omega') \in \mathcal{N},$$

since $\mathcal{N}$ is locally measure convex. This implies clearly $a_t(D) \equiv 1$. Therefore, ρ is acceptance consistent. □

A.8 Proof of Theorem 4.5.

We have already proven one direction in Theorem 4.4. Thus, we only need to show that 'consistency' implies 'measure convexity'. We will focus on the case of acceptance consistency. The case of rejection consistency works analogously.

Let ρ be an M-invariant dynamic risk measure, and let $\mathcal{N}$ be the corresponding acceptance set of the representing static risk measure. Observe that $\mathcal{N}$ is measurable by definition of the functions H_t. Let $c \in \mathbb{R}$ be given, and let γ be a probability measure on $\mathcal{N} \cap \mathcal{M}([-c, c])$. Let $Z \sim unif(0, 1)$ be a random variable independent of $\mathcal{F}_{T-1}$, and let $U \sim unif(0, 1)$ be a $\mathcal{F}_{T-1}$- measurable random variable independent of $\mathcal{F}_{T-2}$. By Borel's theorem[10] there exists a measurable function $\mu : [0, 1] \to \mathcal{N}$ such that $\mu(U) \sim \gamma$. We define a kernel from $\mathcal{M}_1(\mathbb{R})$ to $\mathbb{R}$ by

$$\begin{cases} \mathcal{M}_1(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) & \to \quad [0, 1] \\ (\nu, A) & \mapsto \quad \nu(A) \end{cases}$$

By the kernel randomization lemma[11] there exists a measurable function

$$q : \mathcal{M}_1(\mathbb{R}) \times [0, 1] \to \mathbb{R}$$

[10] See Theorem 2.19 in Kallenberg (1997).
[11] See Lemma 2.22 in Kallenberg (1997).

such that $q_\nu(Z) = q(\nu, Z) \sim \nu$. Clearly, the composite function $q_{\mu(\cdot)}(\cdot) : [0,1]^2 \to \mathbb{R}$ is measurable. We define the random variable $K := q_{\mu(U)}(Z) \in [-c, c]$, and the financial position $D := K \cdot e_T \in \mathcal{D}$. We obtain that for all $\omega \in \Omega$,

$$\mathcal{L}(K|\mathcal{F}_{T-1})(\omega) = \mu(U(\omega)) \in \mathcal{N}, \tag{9}$$

$$\mathcal{L}(K|\mathcal{F}_{T-2}) = \mathcal{L}(K) = \int_{\mathcal{N}} \nu\gamma(d\nu). \tag{10}$$

Equation (9) implies $a_{T-1}(D) \equiv 1$. From acceptance consistence follows $a_{T-2}(D) \equiv 1$. Thus, $\int_{\mathcal{N}} \nu\gamma(d\nu) \overset{(10)}{=} \mathcal{L}(K|\mathcal{F}_{T-2}) \in \mathcal{N}$. □

A.9 Proof of Theorem 5.3

(1) ⇒ (2): Choose $x_1 \in \mathbb{R}$ with $\delta_{x_1} \in \mathcal{N}$ according to (5), and let $x_2 \in \mathbb{R}$, $\delta_{x_2} \in \mathcal{N}^c$. Define the function $g : \mathbb{R} \to \mathbb{R}$ as follows: We set $g(x_1) = 0$ and $g(x_2) = 1$. Let $z := \sup\{0 \le \alpha \le 1 : \alpha\delta_{x_2} + (1-\alpha)\delta_{x_1} \in \mathcal{N}\}$. Since $\mathcal{N}$ is ψ-weakly closed, the supremum is actually a maximum. Thus, $z \neq 1$, since $\delta_{x_2} \notin \mathcal{N}$. By (5) $z > 0$, hence $z \in (0,1)$. Hence, z is in the interior of the convex hull of the range of g.

Since $\mathcal{N}$ is ψ-weakly closed, it follows from inverse monotonicity that there exists $r \in \mathbb{R}$ such that $[r, \infty) = \{y \in \mathbb{R} : \ \delta_y \in \mathcal{N}\}$, $(-\infty, r) = \{y \in \mathbb{R} : \ \delta_y \in \mathcal{N}^c\}$.

If $y \in [r, \infty)$, define

$$\alpha(y) := \sup\{0 \le \alpha \le 1 : \alpha\delta_{x_2} + (1-\alpha)\delta_y \in \mathcal{N}\}.$$

Since $\mathcal{N}$ is ψ-weakly closed, the supremum is actually a maximum. Thus $\alpha(y) \neq 1$, since $\delta_{x_2} \notin \mathcal{N}$. Hence, $1 - \alpha(y) \neq 0$, and we may define

$$g(y) := \frac{z - \alpha(y)}{1 - \alpha(y)}.$$

Inverse monotonicity implies additionally that $y \mapsto \alpha(y)$ is increasing on $[r, \infty)$. Hence, $y \mapsto g(y) = 1 + \frac{z-1}{1-\alpha(y)}$ is decreasing on $[r, \infty)$, since $z - 1 < 0$.

If $y \in (-\infty, r)$, define

$$\alpha(y) := \sup\{0 \le \alpha \le 1 : \alpha\delta_y + (1-\alpha)\delta_{x_1} \in \mathcal{N}\}.$$

Observe that $\alpha(y) \neq 1$, since $\delta_y \notin \mathcal{N}$. By (5) we have $\alpha(y) \neq 0$. We let

$$g(y) := \frac{z}{\alpha(y)}.$$

Inverse monotonicity implies that $y \mapsto \alpha(y)$ is increasing on $(-\infty, r)$. Hence $y \mapsto g(y)$ is decreasing on $(-\infty, r)$.

Moreover, note that on the one hand $g(y) \ge z$ for $y \in (-\infty, r)$. On the other hand, $g(y) = \frac{z}{1-\alpha(y)} - \frac{\alpha(y)}{1-\alpha(y)} \le z$ for $y \in [r, \infty)$. Hence, $g : \mathbb{R} \to \mathbb{R}$ is a decreasing function. We set $l(-x) = g(x)$. For simple probability measures $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$ of the form

$$\mu = \sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i},$$

$\alpha_i \geq 0$, $x_i \in \mathbb{R}$ $(i = 1, 2, \ldots, n)$, $\sum_{i=1}^{n} \alpha_i = 1$, $n \in \mathbb{N}$, we will show that

$$\mu \in \mathcal{N} \Leftrightarrow \int g(x)\mu(dx) \leq z.$$

Let $\mu = \sum_{i=1}^{n} \alpha_i \cdot \delta_{x_i}$ be given. We denote by $\mathcal{M}$ the convex hull of $\{\delta_{x_i} : \ i = 1, 2, \ldots, n\}$. The simplex $\mathcal{M}$ is a convex subset of the n-dimensional vector space spanned by $\{\delta_{x_i} : \ i = 1, 2, \ldots, n\}$. Let $\mathcal{A} := \mathcal{N} \cap \mathcal{M}$, $\mathcal{B} = \mathcal{N}^c \cap \mathcal{M}$. Then $\mathcal{M} = \mathcal{A} \cup \mathcal{B}$, $\mathcal{A} \cap \mathcal{B} = \emptyset$, the sets $\mathcal{A}$ and $\mathcal{B}$ are both convex, and $\mathcal{A}$ is closed in the Euclidian topology. We can therefore find an affine functional $h : \mathcal{M} \to \mathbb{R}$ and $q \in \mathbb{R}$ such that

$$\begin{aligned} h(\mu) &\leq q, \qquad \mu \in \mathcal{A}, \\ h(\mu) &> q, \qquad \mu \in \mathcal{B}. \end{aligned}$$

We define

$$k := \frac{h - h(\delta_{x_1})}{h(\delta_{x_2}) - h(\delta_{x_1})}.$$

Then

$$\begin{aligned} k(\mu) &\leq \frac{q - h(\delta_{x_1})}{h(\delta_{x_2}) - h(\delta_{x_1})}, \qquad \mu \in \mathcal{A}, \\ k(\mu) &> \frac{q - h(\delta_{x_1})}{h(\delta_{x_2}) - h(\delta_{x_1})}, \qquad \mu \in \mathcal{B}. \end{aligned}$$

We show now that $g(x_i) = k(\delta_{x_i})$. For $i = 1, 2$ the claim is immediate from the definition of k. This implies that

$$k(\alpha\delta_{x_2} + (1-\alpha)\delta_{x_1}) = \alpha.$$

Hence,

$$\begin{aligned} z &= \sup\{0 \leq \alpha \leq 1 : \ \alpha\delta_{x_2} + (1-\alpha)\delta_{x_2} \in \mathcal{N}\} \\ &= \sup\left\{0 \leq \alpha \leq 1 : \ \alpha \leq \frac{q - h(\delta_{x_1})}{h(\delta_{x_2}) - h(\delta_{x_1})}\right\} = \frac{q - h(\delta_{x_1})}{h(\delta_{x_2}) - h(\delta_{x_1})}. \end{aligned}$$

Let now $i \neq 1, 2$. Assume first that $x_i \in [r, \infty)$. This implies that

$$\begin{aligned} \alpha(x_i) &= \sup\{0 \leq \alpha \leq 1 : \ \alpha\delta_{x_2} + (1-\alpha)\delta_{x_i} \in \mathcal{N}\} \\ &= \sup\{0 \leq \alpha \leq 1 : \ \alpha + (1-\alpha)k(\delta_{x_i}) \leq z\}. \end{aligned}$$

Observe that $\alpha(x_i) \neq 1$ and that $\alpha \mapsto \alpha + (1-\alpha)k(\delta_{x_i})$ is continuous. Hence, the last equation is satisfied, if and only if $\alpha(x_i) + (1-\alpha(x_i))k(\delta_{x_i}) = z$, i.e.

$$k(\delta_{x_i}) = \frac{z - \alpha(x_i)}{1 - \alpha(x_i)} = g(x_i).$$

Second, consider the case $x_i \in (-\infty, r)$. Then

$$\begin{aligned} \alpha(x_i) &= \sup\{0 \leq \alpha \leq 1 : \alpha\delta_{x_i} + (1-\alpha)\delta_{x_1} \in \mathcal{N}\} \\ &= \sup\{0 \leq \alpha \leq 1 : \alpha k(\delta_{x_i}) \leq z\}. \end{aligned}$$

Observe that $\alpha(x_i) \neq 1$ and that $\alpha \mapsto \alpha k(\delta_{x_i})$ is continuous. Hence, the last equation is satisfied, if and only if $\alpha(x_i)k(\delta_{x_i}) = z$, i.e.

$$k(\delta_{x_i}) = \frac{z}{\alpha(x_i)} = g(x_i).$$

Finally, we obtain for $\mu = \sum_{i=1}^{n} \alpha_i \delta_{x_i}$ that

$$\mu \in \mathcal{N} \Leftrightarrow k(\mu) \leq z \Leftrightarrow \sum_{i=1}^{n} \alpha_i g(x_i) \leq z \Leftrightarrow \int g(x)\mu(dx) \leq z.$$

Next we prove that g is right-continuous, thus l left-continuous. Since g is decreasing, $g(x+)$ exists for each $x \in \mathbb{R}$. We have already shown that $g(x_1) < z$, $g(x_2) > z$. This implies that for given $x \in \mathbb{R}$ we can find $\alpha \in (0, 1]$ and $w \in \mathbb{R}$ such that

$$\alpha g(x+) + (1-\alpha)g(w) = z.$$

Let $x_n \searrow x$. Since g is decreasing, we obtain $\alpha\delta_{x_n} + (1-\alpha)\delta_w \in \mathcal{N}$ $(n \in \mathbb{N})$. Moreover, $\alpha\delta_{x_n} + (1-\alpha)\delta_w$ converges ψ-weakly to $\alpha\delta_x + (1-\alpha)\delta_w$. It follows that $\alpha\delta_x + (1-\alpha)\delta_w \in \mathcal{N}$, since $\mathcal{N}$ is ψ-weakly closed. Thus,

$$z \geq \alpha g(x) + (1-\alpha)g(w) \geq \alpha g(x+) + (1-\alpha)g(w) = z.$$

Therefore, $g(x) = g(x+)$.

Finally, we will show that the representation of $\mathcal{N}$ via the function g is not restricted to simple probability measures. Let $\mu \in \mathcal{N}$. Then there exists a decreasing sequence of simple probability measures $(\mu_n)_n \subseteq \mathcal{M}_{1,c}(\mathbb{R})$ converging to μ ψ-weakly from above. By inverse monotonicity, $(\mu_n)_n \subseteq \mathcal{N}$, thus

$$z \geq \int g(x)\mu_n(dx) \rightarrow \int g(x)\mu(dx).$$

The convergence of the integrals follows from the right-continuity of g.[12] Conversely, let $z \geq \int g(x)\mu(dx)$. Then there exists a decreasing sequence of simple probability

[12] This fact can easily be proven using Skorohod representation and Lebesgue's dominated convergence, since g is bounded on a superset of the supports of the measures $(\mu_n)_n$ and μ.

measures $(\mu_n)_n \subseteq \mathcal{M}_{1,c}(\mathbb{R})$ converging ψ-weakly to μ from above. Since g is decreasing, we obtain $z \geq \int g(x)\mu_n(dx)$, thus $(\mu_n)_n \subseteq \mathcal{N}$. Since $\mathcal{N}$ is ψ-weakly closed, we obtain $\mu \in \mathcal{N}$.
(2) $\Rightarrow$ (1): The convexity of the acceptance and rejection sets of Θ is immediate. We need to show that the acceptance set is ψ-weakly closed.

Let $\psi \in C(\mathbb{R})$, $\psi \geq |g| + 1$ with $g(x) = l(-x)$ $(x \in \mathbb{R})$. We show that the functional $\mu \mapsto \int g(x)\mu(dx)$ is lower semicontinuous with respect to the ψ-weak topology. Since the ψ-weak topology on $\mathcal{M}_{1,c}(\mathbb{R})$ is metrizable, we employ the sequential characterization of lower semicontinuity. Let $z \in \mathbb{R}$ be given, and let $(\mu_n)_n \subseteq \mathcal{M}_{1,c}(\mathbb{R})$, $\mu_n \to \mu \in \mathcal{M}_{1,c}(\mathbb{R})$ ψ-weakly, where $\int g(x)\mu_n(dx) \leq z$ for $n \in \mathbb{N}$.

By Skorohod representation we can find bounded random variables $(X_n)_n$, X on some probability space $(\Omega, \mathcal{F}, P)$ such that $X_n \sim \mu_n$ $(n \in \mathbb{N})$, $X \sim \mu$, $X_n \to X$ P-a.s.

We have $\lim \psi(X_n) = \psi(X)$ P-almost surely, and $\lim \int \psi(X_n)dP = \int \psi(X)dP$. Observe that $\psi(X_n) + g(X_n) \geq 0$ $(n \in \mathbb{N})$. By Fatou's Lemma we obtain that

$$\begin{aligned} & \int \psi(X)dP + z \ \geq \ \int \psi(X)dP + \liminf_n \int g(X_n)dP \\ = \ & \liminf_n \int (\psi(X_n) + g(X_n))dP \ \geq \ \int \liminf_n (\psi(X_n) + g(X_n))dP) \\ = \ & \int \psi(X)dP + \int \liminf_n g(X_n)dP \ \geq \ \int \psi(X)dP + \int g(X)dP. \end{aligned}$$

The last inequality follows from the fact that g is decreasing and right-continuous, since $X_n \to X$ P-almost surely. Hence,

$$z \ \geq \ \int g(X)dP \ = \ \int g(x)\mu(dx).$$

□

Lemma A.7. *Let $I : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ be an affine, ψ-weakly continuous functional. Then there exists $g \in C^{\psi}$ such that*

$$I(\mu) = \int g(x)\mu(dx) \qquad\qquad (\mu \in \mathcal{M}_{1,c}(\mathbb{R})).$$

Proof.
Define $g(x) = I(\delta_x)$. If $x_n \to x$, then $\delta_{x_n} \to \delta_x$ ψ-weakly, hence $g(x_n) \to g(x)$. This implies that g is continuous.
Suppose that g/ψ is unbounded, say

$$\sup_{x \in \mathbb{R}} \frac{g(x)}{\psi(x)} = \infty.$$

Let $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ be a sequence of real numbers such that $g(x_n) \cdot (\psi(x_n))^{-1} \geq n^2$, and let

$$\mu_n \ = \ \left(1 - \frac{1}{n\psi(x_n)}\right) \cdot \delta_0 \ + \ \frac{1}{n\psi(x_n)} \cdot \delta_{x_n}.$$

Then $\mu_n \to \delta_0$ ψ-weakly, but

$$I(\mu_n) \;=\; \left(1 - \frac{1}{n\psi(x_n)}\right) g(0) \;+\; \frac{1}{n} \cdot \frac{g(x_n)}{\psi(x_n)}$$

diverges. Hence, we obtain $g \in C^\psi$.
Finally, we have to show that for $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$,

$$I(\mu) = \int g(x)\mu(dx).$$

The equality does certainly hold for simple probability measures which form a dense subset of $(\mathcal{M}_{1,c}(\mathbb{R}), \tau_\psi)$. Here, τ_ψ denotes the ψ-weak topology. Let now $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$ be arbitrary, and let $(\mu_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}_{1,c}(\mathbb{R})$ be a sequence of simple probability measures converging ψ-weakly to μ. By continuity of I we get, $I(\mu_n) \to I(\mu)$. Since $g \in C^\psi$, we obtain that

$$I(\mu_n) = \int g(x)\mu_n(dx) \to \int g(x)\mu(dx).$$

□

A.10 Counterexample for $AVaR_\lambda$

The acceptance set of $AVaR_\lambda$ ($\lambda \in (0,1)$) is not convex as subsets of the space of probability measures. For each $\lambda \in (0,1)$ this can be demonstrated by the following counterexample.

We let $\mu = \lambda \cdot \text{unif}\,[-1,1] \;+\; (1-\lambda) \cdot \text{unif}\,[1,2]$, $\nu = \delta_0$. Then

$$q_\mu(\gamma) = \frac{2\gamma}{\lambda} - 1, \qquad (\gamma \le \lambda).$$

Hence, $AVaR_\lambda(\mu) = 0$. Moreover, $AVaR_\lambda(\nu) = 0$. This implies $\mu, \nu \in \mathcal{N}$. Let $\alpha = \lambda + \frac{1-\lambda}{2}$. Then $q_{\alpha\nu+(1-\alpha)\mu}(\lambda) = 0$. But

$$q_{\alpha\nu+(1-\alpha)\mu}(\gamma) = \frac{2\gamma}{(1-\alpha)\lambda} - 1, \qquad \left(\gamma \le \frac{(1-\alpha)\lambda}{2}\right).$$

Hence, $AVaR_\lambda(\alpha\nu + (1-\alpha)\mu) > 0$. This implies that $\alpha\nu + (1-\alpha)\mu \notin \mathcal{N}$. The acceptance set of $AVaR_\lambda$ is therefore not a convex subset of the space of probability measures.

A.11 Proof of Corollary 5.9

If l is convex, the corresponding risk measure is clearly convex. We only have to prove the other direction. Assume thus that l is not convex. Then g is not convex, and we can find $x, y \in \mathbb{R}$, $x < y$, such that

$$\frac{g(x) + g(y)}{2} \;<\; g\left(\frac{x+y}{2}\right).$$

Because z is in the interior of the convex hull of the range of g, we can always find $w \in \mathbb{R}$ and $\alpha \in [0,1)$ such that

$$\alpha g(w) \; + \; (1-\alpha) \cdot \left(\frac{g(x)+g(y)}{2} \right) \; \leq \; z \; < \; \alpha g(w) \; + \; (1-\alpha) \cdot g\left(\frac{x+y}{2} \right).$$

We define the following random variables on $((0,1), \lambda)$:

$$\begin{aligned} Z_1 &= w \cdot \mathbf{1}_{(0,\alpha)} \; + \; x \cdot \mathbf{1}_{[\alpha, 1-\alpha/2)} \; + \; y \cdot \mathbf{1}_{[1-\alpha/2,1)} \\ Z_2 &= w \cdot \mathbf{1}_{(0,\alpha)} \; + \; y \cdot \mathbf{1}_{[\alpha, 1-\alpha/2)} \; + \; x \cdot \mathbf{1}_{[1-\alpha/2,1)} \end{aligned}$$

Then Z_1 and Z_2 are both acceptable, since for $i = 1,2$,

$$\int g(Z_i) d\lambda \; = \; \alpha g(w) \; + \; (1-\alpha) \left(\frac{g(x)+g(y)}{2} \right) \; \leq \; z.$$

We define $Z := \frac{Z_1+Z_2}{2} = w \cdot \mathbf{1}_{(0,\alpha)} + \frac{x+y}{2} \cdot \mathbf{1}_{[1-\alpha,1)}$, and obtain

$$\int g(Z) d\lambda \; = \; \alpha g(w) \; + \; (1-\alpha) \cdot g\left(\frac{x+y}{2} \right) \; > \; z.$$

Hence, Z is not acceptable, contradicting the convexity of Θ.

□

A.12 Proof of Lemma 5.10

We apply Theorem 4.61 of Föllmer & Schied (2002c). For $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$, let $P := \mu$ and $X := id$. By $\mathcal{X}$ we denote the class of all bounded measurable functions. Of course, $(\mathbb{R}, \mathcal{B}, P)$ is not necessarily atomless. Nevertheless, if $\mathcal{L}(Y)$ $(Y \in \mathcal{X})$ denotes the distribution of Y under P, then $\rho(Y) := \Theta(\mathcal{L}(Y))$ $(Y \in \mathcal{X})$ defines a convex risk measure on $\mathcal{X}$ which satisfies the conditions of Proposition 4.59 and Theorem 4.61 of Föllmer & Schied (2002c). This implies Lemma 5.10.

□

A.13 Proof of Corollary 5.13

First, let $l(x) = z + \alpha x^+ - \beta x^-$ be given. Since $\alpha \geq \beta > 0$, the loss function l is convex. Hence, l induces a convex risk measure. Let $\mu \in \mathcal{N}$, and let $X \sim \mu$ be a random variable on some atomless probability space $(\Omega, \mathcal{F}, P)$. Then for $\lambda \geq 0$,

$$\int l(-\lambda X) dP = z + \lambda \int (l(-X) - z) dP \leq z.$$

This implies that $\mathcal{L}(\lambda X) \in \mathcal{N}$. Hence, Θ is positively homogeneous.

Conversely, let Θ be a coherent risk measure that satisfies the hypotheses. Then Θ can be represented by a continuous and convex loss function l and a threshold level $z \in \mathbb{R}$ in the interior of the range of l. Since Θ is positively homogeneous, $\delta_y \in \mathcal{N}$ for $y \in [0, \infty)$ and $\delta_y \in \mathcal{N}^c$ for $y \in (-\infty, 0)$. This implies that $l(0) = z$. Subtracting z, we may w.l.o.g. assume that $z = 0$ and $l(0) = 0$. Let $g(x) := l(-x)$.

Suppose that there exist $x' \in \mathbb{R}$, $\lambda' \geq 0$ such that $g(\lambda' x') \neq \lambda' g(x')$. Since g is convex and $g(0) = 0$, this implies that there exist $x \in \mathbb{R}$ and $\lambda > 1$ such that $g(\lambda x) > \lambda g(x)$. Since $z = 0$ lies in the interior of the range of g, we can find $w_1, w_2 \in \mathbb{R}$ such that $g(w_1) < 0 < g(w_2)$. Therefore there exist $w \in \mathbb{R}$ and $\alpha \in (0, 1]$ such that

$$\alpha g(x) + (1 - \alpha) g(w) = 0.$$

Hence, $\alpha \delta_x + (1 - \alpha)\delta_w \in \mathcal{N}$. Since g is convex with $g(0) = 0$, $g(\lambda w) \geq \lambda g(w)$. Since $g(\lambda x) > \lambda g(x)$, we obtain

$$\alpha g(\lambda x) + (1 - \alpha) g(\lambda w) > 0.$$

This implies that $\alpha \delta_{\lambda x} + (1 - \alpha)\delta_{\lambda w} \notin \mathcal{N}$ – contradicting the assumption of coherence. Altogether we obtain that for $x \in \mathbb{R}$, $\lambda \geq 0$ it holds that $\lambda g(x) = g(\lambda x)$. This implies that g is of the form

$$g(x) = \alpha x^- - \beta x^+$$

for $\alpha, \beta \in \mathbb{R}$. $\alpha, \beta \geq 0$, since g is decreasing. The inequality $\alpha \geq \beta$ follows from the convexity of g. Finally, $\alpha, \beta > 0$, because 0 lies in the interior of the range of g. □

A.14 Proof of Theorem 5.14

By Corollary 4.2 there exists a unique risk measure Θ such that ρ can be represented according to (6). By Theorem 4.5 the acceptance set $\mathcal{N}$ and the rejection set $\mathcal{N}^c$ are locally measure convex, thus convex. Hence, $\mathcal{N}$ can be represented according to Theorem 5.3 for some loss function $l : \mathbb{R} \to \mathbb{R}$. The convexity of ρ implies the convexity of Θ. This implies by Corollary 5.9 that l is convex and therefore continuous. Hence, Θ is the shortfall risk measure associated with the continuous and convex loss function l.

□

References

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber & David Heath (1999), 'Coherent measures of risk', *Mathematical Finance* **9**(3), 203–228.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, David Heath & Hyehin Ku (2003), Coherent multiperiod risk adjusted values and Bellman's principle. Working paper, ETH Zürich.

Carlier, Guillaume & Rose-Anne Dana (2003), 'Core of convex distortions of a probability', *Journal of Economic Theory* **113**, 199–222.

Cheridito, Patrick, Freddy Delbaen & Michael Kupper (2003), Coherent and convex risk meaures for càdlàg processes. Working paper, ETH Zürich.

Delbaen, Freddy (2000), Coherent risk measures. Lecture Notes, Scuola Normale Superiore di Pisa.

Delbaen, Freddy (2002), Coherent risk measures on general probability spaces, *in* K. Sandmann & P. J. Schönbucher, eds, 'Advances in Finance and Stochastics', Springer-Verlag Berlin, pp. 1–38.

Delbaen, Freddy (2003), The structure of m-stable sets and in particular of the set of risk-neutral measures. Working paper, ETH Zürich.

Detlefsen, Kai (2003), Bedingte und mehrperiodische Risikomaße. Diplomarbeit, Humboldt-Universität zu Berlin.

Epstein, Larry G. & Martin Schneider (2003), 'Recursive multiple-priors', *Journal of Economic Theory* **113**, 1–31.

Föllmer, Hans & Alexander Schied (2002*a*), 'Convex measures of risk and trading constraints', *Finance and Stochastics* **6**(4), 429–448.

Föllmer, Hans & Alexander Schied (2002*b*), Robust preferences and convex meaures of risk, *in* K. Sandmann & P. J. Schönbucher, eds, 'Advances in Finance and Stochastics', Springer-Verlag Berlin, pp. 39–56.

Föllmer, Hans & Alexander Schied (2002*c*), *Stochastic Finance - An Introduction in Discrete Time*, Walter de Gruyter, Berlin.

Frittelli, Marco & Gianin E. Rosazza (2002), 'Putting order in risk measures', *Journal of Banking and Finance* **26**(7), 1473–1486.

Jaschke, Stefan & Uwe Küchler (2001), 'Coherent risk measures and good-deal bounds', *Finance and Stochastics* **5**(2), 181–200.

Kallenberg, Olav (1997), *Foundations of Modern Probability*, Springer, New York.

Kunze, Mathias (2003), Verteilungsinvarinate konvexe Risikomaße. Diplomarbeit, Humboldt-Universität zu Berlin.

Kusuoka, Shigeo (2001), 'On law invariant coherent risk measures', *Advances in Mathematical Economics* **3**, 83–95.

Owen, Arthur B. (1987), Nonparametric conditional estimation. Ph.D. Thesis, Stanford University.

Rachev, Svetlozar T. (1991), *Probability Metrics and the Stability of Stochastic Models*, John Wiley & Sons, Chichester, New York.

Riedel, Frank (2002), Dynamic coherent risk measures. Working Paper, Stanford University.

Scandolo, Giacomo (2003), Risk measures in a dynamic setting. Ph.D. Thesis, Università degli Studi Milano & Università di Firenze.

Wang, Tan (1996), A characterization of dynamic risk measures. Working paper, U.B.C.

Wang, Tan (1999), A class of dynamic risk measures. Working paper, U.B.C.

Winkler, Gerhard (1985), *Choquet order and simplices with applications to probabilistic models*, Springer-Verlag, Berlin.