Abstract

In this paper we study time-consistent risk measures for returns that are given by a GARCH(1, 1) model. We present a construction of risk measures based on their static counterparts that overcomes the lack of time-consistency. We then study in detail our construction for the risk measures Value-at-Risk (VaR) and Average Value-at-Risk (AVaR). While in the VaR case we can derive an analytical formula for its time-consistent counterpart, in the AVaR case we derive lower and upper bounds to its time-consistent version. Furthermore, we incorporate techniques from Extreme Value Theory (EVT) to allow for a more tail-gared analysis of the corresponding risk measures. We conclude with an application of our results to stock prices to investigate the applicability of our results.

1 Introduction

In the wake of the financial crisis risk management constitutes a constant active field that attracts both mathematical research and quantitative requirements for the practical implementation. Most financial institutions need to abide with the Basel II/III accords that prescribe certain risk management rules to be applied to internal risk control and that are under periodic regulatory supervision. Over the last two decades the key notion of risk management arose in the form of a risk measure referred to as Value-at-Risk (VaR). Simply put, VaR determines the risk capital of a financial institution as the quantile of a profit-and-loss distribution with respect to some prescribed (either by regulation or by internal rules) time horizon and confidence level. An axiomatic approach to the field of risk measures is given by Artzner et al. (1999) in which the notion of the coherent risk measure is introduced and where it has been realized that VaR does not always satisfy the property of coherence. Artzner et al. (1999) introduce a risk measure that amends the lack of coherence that is nowadays known as the Average-Value-at-Risk (AVaR). Both VaR and AVaR are risk measures whose input is the distribution of the profit-and-loss positions and as such fall into the class of law-invariant risk measures. An extension to convex risk measures is given in Föllmer and Schied (2002), which integrates existing notions of risk into the mathematical framework of convex dual theory and, hence,
allows for deep and powerful dual characterizations. In order to account for the dynamic stochastic
evolution of profit-and-loss positions the static risk measurement has been extended to the class of
dynamic risk measures, which treats the risk measure not only as a (nonlinear) expectation but as a
stochastic process, see e.g. Detlefsen and Scandolo (2005) and Riedel (2004) for the extension to the
dynamic setting by means of convex dual theory. It has been realized in this dynamic framework that
most existing static risk measures do not transfer in a straightforward manner into processes without
entailing the spurious effect of violating time-consistency. A time-consistent dynamic risk measure
secures the consistent behavior of a risk measure that, if a portfolio is riskier than another portfolio
at some future time, then this portfolio has been riskier that the other portfolio at any time before.
The literature on time-consistency of risk measures is diverse and rich as different mathematical
viewpoints can be adopted to prevent the consistency property. An incomplete chronicle of research
done in the field of time-consistent risk measures includes Peng (2004), Riedel (2004), Detlefsen and
Scandolo (2005), Penner (2007), Föllmer and Penner (2006), Bion-Nadal (2009). A major result from
the research on time-consistency reveals that in the class of law-invariant risk measures there is only
one risk measure that, upon transfer into a time-dynamic process setting, supports time-consistency,
namely the entropic risk measure (cf. Föllmer and Knispel (2011)).

In parallel to the aforementioned theoretical work statistical models and methods have been developed
to calibrate and integrate risk measures to real world data. As the industry standard VaR and its
coherent counterpart AVaR are law-invariant risk measures, the main goal for an implementation of
(A)VaR is to find a good estimate of the profit-and-loss distribution in the relevant region. In this
field, the major class of estimation methods comprise the historical simulation method, methods based
on Gaussian distribution assumptions and methods based on Extreme Value Theory (EVT). We refer
the reader to McNeil et al. (2005), in particular Chapter 2 and Chapter 7, for a detailed account and
references to methods of profit-and-loss distribution estimation. More background on extreme value
theory can be found in the monograph Embrechts et al. (1997). McNeil and Frey (2000) propose an
implementation of VaR and AVaR that is based on an estimation of the log-returns distribution using
a combination of the EVT approach and a GARCH(1,1) model. Their method proceeds in a two-
step scheme: first, the GARCH(1,1) model mimics the inherent stochastic volatility of financial time
series. Second, rather than fitting the entire innovation process of the GARCH(1,1), they adopt a
Peaks-over-Threshold (POT) approach to the residuals and only consider those residuals that exceed
a critical value. This is in accord with the typically high confidence levels that are imposed on
(A)VaR to zoom into the extreme branch of losses. Applying the POT method to the residuals
rather than directly to the log-returns has the advantage that the fitting procedure to the extremes
only needs to be applied once due to the white noise property of the residuals. By a key result from
Balkema and de Haan (1974) and Pickands (1975) (cf. Embrechts et al. (1997), Section 3.4 and
Section 6.5), the POT residuals are fitted to a Generalized Pareto distribution (GPD) by means of
a pseudo maximum likelihood method. Using these two steps McNeil and Frey (2000) succeed to
estimate (A)VaR by fitting a distribution that adequately accounts for the extremes in the tail and
under mild conditions allows for closed form formulas for VaR and AVaR. However, the underlying
GARCH(1,1) model prevents to obtain closed form formulas for aggregated losses over time. In fact,
McNeil and Frey (2000) have to resort to bootstrapping and Monte Carlo methods to estimate VaR
and AVaR for losses aggregated over e.g. two weeks. In fact, McNeil and Frey (2000) apply VaR
and AVaR conditional on the information at some time point $t$ (represented by the $\sigma$-field $\mathcal{F}_t$) to
risk positions that evolve after $t$. Hence the time point $t$ is fixed and thus the risk assessment fixes
the time of evaluation $t$ and the risk position at some time point after $t$. As such VaR and AVaR
measurements are applied as static risk measures, since the dynamic evolution of VaR and AVaR at
time points after $t$ are neglected. The issue of time-consistency moved into the focus of research after
McNeil and Frey (2000) has been published.
The goal of our paper is to extend the work of McNeil and Frey (2000) to incorporate dynamic time-consistency for VaR and AVaR. We investigate the extension of static risk measures to dynamic counterparts that satisfy time-consistency. A key property to succeed in this transfer is the dynamic programming principle, see Cheridito and Stadje (2009), Cheridito and Kupper (2011). We emphasize that the interpretation for the dynamic time-consistent (A)VaR differs to the interpretation of the static (A)VaR as the dynamic (A)VaR evolves via the composition of the static (A)VaR over time. Composing (A)VaR, however, results in a more conservative risk measurement as the risky positions that are due far in the future not only enter the risk assessment through their own dynamics at the future maturity but rather enter through their risk assessment along any time point up to maturity. Thus risky effects that arise until maturity are cushioned at any time. As a trade-off typical confidence levels like 99% or 99.9% that are requested for static risk measures are not necessary, but can be decreased to a bandwidth of 95%–99%. The two-step estimation scheme from McNeil and Frey (2000) using GARCH (1, 1) and EVT allows us to derive a closed form expression for the dynamic time-consistent VaR that is easily implemented using the estimated GPD and the GARCH parameters. For AVaR however, such a closed form expression cannot be obtained. We can derive closed form lower and upper approximations to AVaR and investigate their accuracy in the context of a practical example. On top of being more conservative than their static counterparts, the dynamic time-consistent risk measures offer the benefit that the risk measurement of aggregated losses, which in e.g. McNeil and Frey (2000) have to be estimated by simulation methods, can now be estimated in a (semi)-closed way by simply aggregating the risk measures of the single positions at different future time points. Although this linearization property is an simple consequence of the dynamic programming principle, it allows for a fast and easy risk assessment of aggregated losses over e.g. two weeks without having to rely on simulation methods.

The paper is structured as follows. In Section 2 we present preliminaries on dynamic risk measures along with the dynamic programming principle characterization and the linearization property for aggregated losses. Moreover, we introduce the GARCH(1, 1) loss model, which establishes the model framework for the entire paper. In Section 3 we apply the new methodology from the previous section to derive a closed form expression for the time-consistent VaR and investigate its properties concerning the evolution over time and the linearization of aggregated losses. Section 4 is devoted to the study of AVaR. Since a closed form expression for time-consistent AVaR is not possible, as an alternative, we derive closed form expressions for pragmatic lower and upper bounds to AVaR and study the properties as in the previous section. The proofs of the results of Sections 3 and 4 are postponed to the Appendix. In the last Section 5 we give a rehash on the part of extreme value theory that is relevant for our purpose, and proceed to testing our results on a data set of financial stock prices.

2 Conditional risk measures

Given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) we consider a filtration \((\mathcal{F}_t)_{t=0}^{T}\) where \(T \in \mathbb{N}\). We denote by \(L^0(\mathcal{F}_t)\) with \(t \in \{0, \ldots, T\}\) the set of all \(\mathcal{F}_t\)-measurable random variables \(X : \Omega \to \mathbb{R}\). In this paper, the space \(L^0(\mathcal{F}_T)\) represents the space of all financial positions for which we need a risk assessment. Typically, we will be interested in losses, i.e. the negatives of log-returns of financial data. Since conditional risk measures are random variables, all properties, equalities and inequalities below hold almost surely with respect to \(\mathbb{P}\), and we assume this throughout without making extra mention of it.

**Definition 2.1.** For \(t \in \{0, \ldots, T\}\) a family of mappings \((\phi_t)_{t=0}^{T}\) with \(\phi_t : L^0(\mathcal{F}_T) \to L^0(\mathcal{F}_t)\) is a dynamic monetary risk measure if it satisfies the following properties:
Corollary 2.4. For all $X,Y \in L^0(\mathcal{F}_T)$ such that $X \geq Y$, for $t = 0, \ldots, T$;

Monotonicity: $\phi_t(X) \geq \phi_t(Y)$ for all $X,Y \in L^0(\mathcal{F}_T)$ such that $X \geq Y$ and $t = 0, \ldots, T$;

Translation invariance: $\phi_t(X + m) = \phi_t(X) + m$ for all $X \in L^0(\mathcal{F}_T)$ and $m \in L^0(\mathcal{F}_t)$, for $t = 0, \ldots, T$.

If $L^0(\mathcal{F}_T)$ represents the space of all profit and loss variables, the above definition leads to the notion of a dynamic monetary utility function, see Definition 2.1 in Cheridito and Kupper (2011). If a dynamic monetary risk measure $\phi$ satisfies in addition to Definition 2.1 (i)-(iii)

- Positive homogeneity: $\phi_t(\lambda X) = \lambda \phi_t(X)$ for all $X \in L^0(\mathcal{F}_T)$ and $\lambda > 0$, for $t = 0, \ldots, T$;
- Subadditivity: $\phi_t(X + Y) \leq \phi_t(X) + \phi_t(Y)$ for all $X,Y \in L^0(\mathcal{F}_T)$, for $t = 0, \ldots, T$,

then we say that $\phi$ is a coherent (dynamic monetary) risk measure.

Definition 2.2. A dynamic monetary risk measure $\phi := (\phi_t)_{t=0}^T$ is time-consistent if

\[ \phi_{t+1}(X) \geq \phi_{t+1}(Y) \text{ implies } \phi_t(X) \geq \phi_t(Y), \]

for all $X,Y \in L^0(\mathcal{F}_T)$, for $t = 0, \ldots, T - 1$.

The following useful characterization of time-consistency can be found in Cheridito and Kupper (2011).

Proposition 2.3. A dynamic monetary risk measure $(\phi_t)_{t=0}^T$ is time-consistent if and only if it satisfies the Bellman principle

\[ \phi_t(X) = \phi_t(\phi_{t+1}(X)) \quad (2.1) \]

for all $X \in L^0(\mathcal{F}_T)$, and $t = 0, \ldots, T - 1$.

It has been noted in Cheridito and Stadje (2009) and Cheridito and Kupper (2011) that there is another way to construct time-consistent dynamic risk measures: let $(\rho_t)_{t=0}^{T-1}$ be an arbitrary dynamic monetary risk measure

\[ \rho_t : L^0(\mathcal{F}_T) \to L^0(\mathcal{F}_t), \quad t = 0, \ldots, T - 1, \]

then the backward iteration

\[ \phi_t(X) := X, \quad \phi_t(X) := \rho_t(\phi_{t+1}(X)), \quad t = 0, \ldots, T - 1, \quad (2.2) \]

defines a process $(\phi_t)_{t=0}^T$ which by definition is a time-consistent dynamic risk measure. The following property is a straightforward consequence of the construction of $(\phi_t)_{t=0}^T$.

Corollary 2.4. For $X \in L^0(\mathcal{F}_T)$ we have for $t = 0, \ldots, T - 1$

\[ \phi_t(X) = \left( \rho_t \circ \rho_{t+1} \circ \cdots \circ \rho_{T-1} \right)(X). \quad (2.3) \]

It is a consequence of the Bellman principle that for losses which aggregate over time the time-consistent risk measure $(\phi_t)_{t=0}^T$ has a linearization property, which allows for a convenient risk assessment of aggregated risk positions on the basis of the single risk components.
Proposition 2.5. For fixed $t \in \{0, \ldots, T\}$ and $m \in \mathbb{N}$ such that $t + m \leq T$ let $X_{t+k} \in L^0(\mathcal{F}_{t+k})$ for $k = 1, \ldots, m$ be given loss positions. We have

$$\phi_t(\sum_{k=1}^{m} X_{t+k}) = \sum_{k=1}^{m} \phi_t(X_{t+k}). \quad (2.4)$$

Proof. According to Corollary 2.4 we have

$$\phi_t(\sum_{k=1}^{m} X_{t+k}) = \left( \rho_t \circ \rho_{t+1} \circ \cdots \circ \rho_{t+m-1} \right) \left( \sum_{k=1}^{m} X_{t+k} \right).$$

By iteratively applying the translation invariance property of $\rho_t$ (see Definition 2.1 (iii)) and the definition of $\phi$ we obtain

$$\phi_t(\sum_{k=1}^{m} X_{t+k}) = \left( \rho_t \circ \rho_{t+1} \circ \cdots \circ \rho_{t+m-2} \right) \left( \sum_{k=1}^{m-1} X_{t+k} + \rho_{t+m-1}(X_{t+m}) \right)$$

$$= \left( \rho_t \circ \rho_{t+1} \circ \cdots \circ \rho_{t+m-3} \right) \left( \sum_{k=1}^{m-2} X_{t+k} + \rho_{t+m-2}(X_{t+m-1}) + \phi_{t+m-2}(X_{t+m}) \right)$$

$$= \left( \rho_t \circ \rho_{t+1} \circ \cdots \circ \rho_{t+m-4} \right) \left( \sum_{k=1}^{m-3} X_{t+k} + \rho_{t+m-3}(X_{t+m-2}) + \phi_{t+m-3}(X_{t+m-1}) + \phi_{t+m-3}(X_{t+m}) \right)$$

$$= \ldots = \sum_{k=1}^{m} \phi_t(X_{t+k}).$$

$\square$

2.1 The GARCH(1,1) model for loss positions

Recall that we are interested in the risk assessment of losses. The focus of this paper is on a particular class of loss processes $(L_t)_{t=1}^{T}$: its dynamics is governed by a GARCH(1,1) process and typically represent (negative) log-returns. It holds that $(L_t)_{t=1}^{T}$ satisfies

$$L_t = \sigma_t Z_t, \quad \sigma_t^2 = a_0 + a_1 L_{t-1}^2 + b\sigma_{t-1}^2, \quad (2.5)$$

where $a_0, a_1, b > 0$ are the model parameters, $\sigma_0$ and $L_0$ are $\mathcal{F}_0$-measurable initial random variables, and $(Z_t)_{t=1}^{T}$ is a strict white noise process (independently identically distributed with zero mean and unit variance). Note also that by (2.5) $\sigma_t$ is measurable with respect to $\mathcal{F}_{t-1}$ for every $t = 1, \ldots, T$.

We denote by $F_Z : \mathbb{R} \to [0, 1]$ and $F_Z^{-1} : [0, 1] \to \mathbb{R}$ the distribution function and the left-continuous quantile function of each $Z_t$, respectively; i.e.,

$$F_Z(z) = \mathbb{P}(Z_t \leq z), \quad F_Z^{-1}(\alpha) = \inf\{x \in \mathbb{R} : F_Z(x) \geq \alpha\}, \quad \alpha \in (0, 1), \quad t = 0, \ldots, T. \quad (2.6)$$

For properties of the quantile function $F_Z^{-1}$ we refer to Resnick [1987], Section 0.2, or Embrechts et al. [1997], Appendix A1.6. We assume that $F_Z$ is strictly increasing, thus $F_Z^{-1}$ is continuous, and that the right endpoint of $Z_t$ is infinite; i.e.,

$$x_F = \inf\{x \in \mathbb{R} : F_Z(x) = 1\} = \infty.$$
Since we often work with distribution tails, we note that $F_Z^{-1}$ can also be represented as

\[ F_Z^{-1}(\alpha) = \inf\{x \in \mathbb{R} : P(Z_t > x) \leq 1 - \alpha\}, \quad \alpha \in (0, 1), \quad t = 0, \ldots, T. \quad (2.7) \]

If necessary we identify $F_Z^{-1}(1)$ with $x_F = \infty$ and $F_Z^{-1}(0) = \sup\{x \in \mathbb{R} : F_Z(x) = 0\} \geq -\infty$. Since $Z$ has infinite right endpoint, and $\alpha$ close to 1, $F_Z^{-1}(\alpha)$ is as a rule positive.

### 3 Conditional time-consistent Value-at-Risk

In this section we study Value-at-Risk (VaR) in the framework of dynamic time-consistent risk measures. One typically considers $L \in L^0(\mathcal{F}_T)$ which represents a large loss position, for which the probability of $L$ exceeding a loss threshold $u > 0$ should be bounded by a small probability $1 - \alpha$, i.e. $\alpha$ is typically close to 1. The smallest loss threshold $u$ which satisfies this bound is the VaR$^\alpha$. Several versions of the (conditional) VaR definition can be found in the literature. In analogy to (2.7) we work throughout with the following, which caters best to the purpose of the treatments in this paper.

**Definition 3.1.** Given a loss position $L \in L^0(\mathcal{F}_T)$ the Value-at-Risk at level $\alpha \in (0, 1)$ at time $t \in \{0, \ldots, T\}$ for $L$ is defined by

\[ \text{VaR}_L^\alpha(t) := \operatorname{essinf}\{m \in L^0(\mathcal{F}_t) : \mathbb{P}(L \leq m \mid \mathcal{F}_t) \geq \alpha\}, \quad (3.1) \]

An immediate consequence of this definition is that $\text{VaR}_L^\alpha$ maps $L^0(\mathcal{F}_T)$ into the positive real half-line.

**Remark 3.2.** If $L$ has infinite right endpoint and $\alpha$ is close to 1, then $\text{VaR}_L^\alpha(L) \in L^0(\mathcal{F}_T)$ where $L^0(\mathcal{F}_t)$ denotes the space of all nonnegative $\mathcal{F}_t$-measurable loss positions. We will assume this throughout the paper. □

#### 3.1 Time-consistent VaR for single day losses

We start this section with the following example, which is the core object of interest in [McNeil and Frey] (2000).

**Example 3.3.** For $t = 0, \ldots, T - 1$ let $L_{t+1}$ be given by (2.5). Then $\text{VaR}_L^\alpha(L_{t+1})$ is the 1-day-ahead VaR, which can be computed straightforwardly as

\[ \text{VaR}_L^\alpha(L_{t+1}) = \operatorname{essinf}\{m \in L^0_+(\mathcal{F}_t) : \mathbb{P}(\sigma_{t+1}Z_{t+1} \leq m \mid \mathcal{F}_t) \geq \alpha\}. \]

Since $\sigma_{t+1}$ is $\mathcal{F}_t$-measurable, we also have that $\tilde{m} := m/\sigma_{t+1}$ is $\mathcal{F}_t$-measurable. Using the independence between $Z_{t+1}$ and $\mathcal{F}_t$, and also (2.7), we can continue

\[ \text{VaR}_L^\alpha(L_{t+1}) = \sigma_{t+1} \operatorname{essinf}\{\tilde{m} \in L^0_+(\mathcal{F}_t) : \mathbb{P}(Z_{t+1} \leq \tilde{m} \mid \mathcal{F}_t) \geq \alpha\} = \sigma_{t+1} \operatorname{inf}\{\tilde{m} \in \mathbb{R}_+: \mathbb{P}(Z_{t+1} \leq \tilde{m}) \geq \alpha\} = \sigma_{t+1} F_{Z_{t+1}}^{-1}(\alpha). \quad (3.2) \]

□

There are examples showing that Value-at-Risk from Definition 3.1 is not time-consistent (e.g. [Cheridito and Stadje] (2009) or [Föllmer and Schied] (2011, Example 11.13)). As the GARCH(1,1) model (2.5) is defined by an iteration, one could hope that for this specific model VaR$^\alpha$ is time-consistent. However, this is not true and we provide a counterexample, which makes use of Proposition 2.3.
Example 3.4. In order to see why in the framework of GARCH(1,1) losses VaR cannot be time-consistent, recall that according to Proposition 2.3 VaR is time-consistent if and only if it satisfies the dynamic programming principle

$$\text{VaR}_t^\alpha = \text{VaR}_t^\alpha \circ \text{VaR}_t^\alpha, \quad t = 0, \ldots, T - 1.$$  

For $t \in \{0, \ldots, T - 2\}$, by (3.2), we have $\text{VaR}_{t+1}^\alpha(L_{t+2}) = \sigma_{t+2}F_Z^{-1}(\alpha)$ and, hence,

$$\text{VaR}_t^\alpha(\text{VaR}_{t+1}^\alpha(L_{t+2})) = \text{VaR}_t^\alpha(\sigma_{t+2}F_Z^{-1}(\alpha)).$$

We compute $\text{VaR}_t^\alpha(L_{t+2})$ and $\text{VaR}_t^\alpha(\sigma_{t+2}F_Z^{-1}(\alpha))$ for the GARCH(1,1) model. For simplicity assume that the white noise variables $(Z_t)_{t=0}^T$ have a density supported on the whole of $\mathbb{R}$. If at least one of $\sigma_0$ or $L_0$ also have a density, then all $(L_t)_{t=0}^T$ and $(\sigma_t)_{t=0}^T$ have densities, and we assume this, so all distribution functions are continuous in the following. Then

$$\text{VaR}_t^\alpha(\sigma_{t+2}F_Z^{-1}(\alpha)) = m^* = \text{essinf}\{m \in L^0(\mathcal{F}_t) : \mathbb{P}(\sigma_{t+2}F_Z^{-1}(\alpha)^2 \leq m^2 \mid \mathcal{F}_t) \geq \alpha\}$$

$$= \text{essinf}\{m \in L^0(\mathcal{F}_t) : \mathbb{P}(a_0 + \sigma_{t+1}Z_{t+1}^2 + b) \leq m^2 F_Z^{-1}(\alpha)^2 \mid \mathcal{F}_t) \geq \alpha\}$$

which leads by $\mathcal{F}_t$-measurability of $\sigma_{t+1}$, independence of $Z_{t+1}$ of $\mathcal{F}_t$, and continuity of the distribution function of $Z_{t+1}$ to

$$m^* = \sqrt{\sigma_{t+1}^2(a_1F_Z^{-1}(\alpha)^2 + b) + a_0} F_Z^{-1}(\alpha)^2. \quad (3.3)$$

Next we compute

$$\text{VaR}_t^\alpha(L_{t+2}) = m^{**} = \text{essinf}\{m \in L^0(\mathcal{F}_t) : \mathbb{P}(\sigma_{t+2}Z_{t+2}^2 \leq m^2 \mid \mathcal{F}_t) \geq \alpha\}$$

$$= \text{essinf}\{m \in L^0(\mathcal{F}_t) : \mathbb{P}(a_0 + \sigma_{t+1}Z_{t+1}^2 + b) Z_{t+2}^2 \leq m^2 \mid \mathcal{F}_t) \geq \alpha\}.$$  

Now, if $m^{**} = m^*$, then the right hand side is equivalent to (note that $F_Z^{-1}(\alpha) = F_Z^{-1}(\alpha)^2$)

$$\alpha = \mathbb{P}\{(a_0 + \sigma_{t+1}^2(a_1Z_{t+1}^2 + b)) Z_{t+2}^2 \leq (a_0 + \sigma_{t+1}^2(a_1F_Z^{-1}(\alpha)^2 + b)) F_Z^{-1}(\alpha)^2 \mid \mathcal{F}_t\}$$

$$= \alpha^2 \mathbb{P}\{(a_0 + \sigma_{t+1}^2(a_1Z_{t+1}^2 + b)) Z_{t+2}^2 \leq (a_0 + \sigma_{t+1}^2(a_1F_Z^{-1}(\alpha)^2 + b)) F_Z^{-1}(\alpha)^2 \mid Z_{t+1}^2 \leq \kappa^2, Z_{t+2}^2 \leq \kappa^2, \mathcal{F}_t\}$$

$$\leq \alpha^2,$$  

(3.4)

which is a contradiction for $\alpha \in (0, 1)$. \hfill \square

Cheridito and Stadje (2009) propose to amend this deficiency of VaR using the backward iteration (2.2). This gives rise to the following definition.

**Definition 3.5.** Given a loss position $L \in L^0(\mathcal{F}_T)$ and $\text{VaR}_0^\alpha$ from Definition 3.1 Then the time-consistent Value-at-Risk at level $\alpha \in (0, 1)$ for $L$ is defined by

$$\text{VaR}_T^\alpha(L) := \text{VaR}_T^\alpha(L) = L, \quad \text{VaR}_t^\alpha(L) := \text{VaR}_t^\alpha(\text{VaR}_{t+1}^\alpha(L)), \quad t = 0, \ldots, T - 1. \quad (3.5)$$

In the notation of the construction from the recursion (2.2), this corresponds to $\rho_t := \text{VaR}_t^\alpha$ and $\phi_t := \text{VaR}_t^\alpha$. The following result is an obvious consequence of the construction of $(\text{VaR}_T^\alpha(L))_{t=0}^T$.

**Corollary 3.6.** For every loss position $L \in L^0(\mathcal{F}_T)$ we have for $t = 0, \ldots, T - 1$

$$\text{VaR}_t^\alpha(L) = \left(\text{VaR}_t^\alpha \circ \text{VaR}_{t+1}^\alpha \circ \cdots \circ \text{VaR}_{T-1}^\alpha\right)(L). \quad (3.6)$$
The choice of the GARCH\((1,1)\) model \((2.5)\) entails the convenient feature that the \(m\)-day ahead VaR assessment allows for a closed form solution. More precisely, we can derive an analytical solution for the time \(t\) risk assessment of the GARCH\((1,1)\) loss at terminal time \(T\) as follows (as usual we set \(\sum_{k=0}^{\infty} a_n = 0\)). The proof is given in Appendix A.

**Theorem 3.7.** Let \(L^T_{s,0}\) be the loss process given by the GARCH\((1,1)\) model \((2.5)\). Let \(\alpha \in (0,1)\) be such that \(F_Z^{-1}(\alpha) > 0\). Then we have

\[
\widehat{\text{VaR}}^\alpha_t(L_T) = F_Z^{-1}(\alpha) \sqrt{\mathcal{P}_t(a_1 F_Z^{-1}(\alpha)^2 + b)}, \quad t = 0, \ldots, T - 1,
\]

where \(\mathcal{P}_t : \mathbb{R} \to \mathbb{R}\) is an \(\mathcal{F}_t\)-measurable mapping given by

\[
\mathcal{P}_t(x) = a_0 \sum_{k=0}^{T-t-2} x^k + \sigma_{t+1}^2 x^{T-t-1}, \quad t = 0, \ldots, T - 1.
\]

The backward recursion giving rise to \((3.7)\) easily allows for computing \(\widehat{\text{VaR}}^\alpha\) for loss positions \(L_T\) at various maturities \(T\) by appropriately scaling the 1-day-ahead \(\text{VaR}^\alpha\).

**Corollary 3.8.** Let the assumptions of Theorem 3.7 be in force. Then we have for fixed \(t \in \{1, \ldots, T - 1\}\) and \(m \in \mathbb{N}\) such that \(t + m \leq T\)

\[
\widehat{\text{VaR}}^\alpha_t(L_{t+m+1}) = F_Z^{-1}(\alpha) \sigma_{t+1} (a_1 F_Z^{-1}(\alpha)^2 + b)^{m/2}.
\]

**Proof.** Define \(x := a_1 F_Z^{-1}(\alpha)^2 + b\). For fixed \(t > 0\) and \(m \in \mathbb{N}\) note that it follows from \((3.7)\)

\[
\frac{\widehat{\text{VaR}}^\alpha_t(L_{t+m+1})}{\widehat{\text{VaR}}^\alpha_t(L_{t+m})} = \left( \frac{a_0 \sum_{k=0}^{m-1} x^k + \sigma_{t+1}^2 x^{m-1}}{a_0 \sum_{k=0}^{m-2} x^k + \sigma_{t+1}^2 x^{m-1}} \right)^{1/2} = \sqrt{x}.
\]

We can hence conclude from

\[
\frac{\widehat{\text{VaR}}^\alpha_t(L_{t+m})}{\widehat{\text{VaR}}^\alpha_t(L_{t+1})} = \prod_{j=1}^{m} \frac{\widehat{\text{VaR}}^\alpha_t(L_{t+j+1})}{\widehat{\text{VaR}}^\alpha_t(L_{t+j})} = x^{m/2},
\]

and the fact that by \((3.7)\), \(\widehat{\text{VaR}}^\alpha_t(L_{t+1}) = \text{VaR}^\alpha_t(L_{t+1}) = F_Z^{-1}(\alpha) \sigma_{t+1}\), the validity of \((3.9)\). \(\square\)

### 3.2 Time-consistent VaR for aggregated losses

We now come to the estimation of the \(m\)-day-ahead \(\text{VaR}^\alpha\). So far, we have considered risk positions at a fixed day \(T\) that is ahead of time \(t < T\) up to which information in the form of the filtration \(\mathcal{F}_t\) is available. The \(m\)-day-ahead \(\text{VaR}^\alpha\) is a risk assessment of aggregated losses \(L_{t+k}\) from the time period \([t+1, t+m]\) for \(t + m \leq T\), i.e.

\[
\rho_t \left( \sum_{k=1}^{m} L_{t+k} \right)
\]

with \(\rho\) denoting a generic time-consistent risk measure. For the time-consistent VaR, it is an immediate consequence of Proposition 2.5 that it linearizes across the aggregation of GARCH\((1,1)\) losses. Though straightforward, we formulate this fact as a result.

**Proposition 3.9.** Under the assumptions of Theorem 3.7 we have for fixed \(t \in \{1, \ldots, T - 1\}\) and \(m \in \mathbb{N}\) such that \(t + m \leq T\),

\[
\widehat{\text{VaR}}^\alpha_t \left( \sum_{k=1}^{m} L_{t+k} \right) = \sum_{k=1}^{m} \widehat{\text{VaR}}^\alpha_t(L_{t+k}).
\]

(3.10)
4 Conditional time-consistent Average Value-at-Risk

This section is devoted to the study of time-consistent alternatives for the Average Value-at-Risk (AVaR). Without going into details, AVaR belongs to the class of coherent risk measures, a property that VaR fails to satisfy for all probability distributions. Due to coherence AVaR is commonly considered as a more reasonable rectification of VaR. More details we refer the reader to Föllmer and Schied (2011, Chapter 4). The following definition relates AVaR to VaR.

**Definition 4.1.** Given a loss position \( L \in L^0(\mathcal{F}_T) \) the Average Value-at-Risk at level \( \alpha \in (0, 1) \) at time \( t \in \{0, \ldots, T\} \) is on its support given by

\[
\text{AVaR}^\alpha_t(L) = \frac{1}{1 - \alpha} \int_0^1 \text{VaR}^u_t(L) \, du. \tag{4.1}
\]

with \( \text{VaR}^\alpha_t(L) \) as in Definition 3.1.

Whereas VaR quantifies the risk associated to one particular level of risk, reflected in the choice of \( \alpha \), AVaR as an integrated VaR takes into account VaR at the entire bandwidth of risk levels between \( \alpha \) and 1 and thus better reflects volume of extreme risks that VaR might neglect.

The following is the analog of a fact well-known for unconditional AVaR (e.g. Lemma 2.16 of McNeil et al. (2005)).

**Remark 4.2.** If the loss position \( L \in L^0(\mathcal{F}_T) \) has a continuous distribution function, then

\[
\text{AVaR}^\alpha_t(L) = \sigma_t \mathbb{E} \left[ Z_t \mid Z_t > F_{\mathcal{F}_t}^{-1}(\alpha) \right], \quad t = 0, \ldots, T. \tag{4.2}
\]

Due to (4.2), AVaR is often also referred to as conditional VaR or Expected Shortfall.

4.1 Time-consistent AVaR for single day losses

We focus again on the GARCH\((1,1)\) model from (2.5).

**Example 4.3.** Assume the setting as in Example 3.3. For \( t = 0, \ldots, T−1 \) let \( L_{t+1} \) be given by (2.5). Then \( \text{AVaR}^\alpha_t(L_{t+1}) \) is the 1-day-ahead-AVaR. If the innovations \( (Z_t)_{t \geq 0} \) have a continuous distribution function \( F_Z \), then a simple computation yields

\[
\text{AVaR}^\alpha_t(L_{t+1}) = \sigma_{t+1} \mathbb{E} \left[ Z_t \mid Z_t > F_Z^{-1}(\alpha) \right], \quad t = 0, \ldots, T,
\]

see e.g. McNeil and Frey (2000).

In analogy to Section 3 a time-consistent version of AVaR is constructed as follows.

**Definition 4.4.** Given a loss position \( L \in L^0(\mathcal{F}_T) \) and the \( \text{AVaR}^\alpha_t(L) \) as in Definition 4.1. Then the time-consistent Average Value-at-Risk at level \( \alpha \in (0, 1) \) at time \( t \in \{0, \ldots, T\} \) is defined by

\[
\bar{\text{AVaR}}^\alpha_T(L) := L, \quad \bar{\text{AVaR}}^\alpha_t(L) := \text{AVaR}^\alpha_t(\text{AVaR}^\alpha_{t+1}(L)), \quad t = 0, \ldots, T−1. \tag{4.3}
\]

For the Average Value-at-Risk of the squared loss \( L_T^2 \) at time \( T \) we can derive an explicit formula similar to (3.7). Note that, though AVaR of \( L_T^2 \) allows for an interpretation as the conditional volatility at time \( T \), our purpose of investigation is to employ AVaR of \( L_T^2 \) to derive pragmatic bounds to AVaR itself, see Section 4.2 below.

We start with a result analog to Theorem 3.7 and recall that \( \sum_{k=0}^{1} = 0 \). The proof is found in Appendix B.
Theorem 4.5. Let \((L_t)_{t=0}^T\) be given by the GARCH(1,1) model \((2.5)\). Then we have for the squared loss \(L_t^2\) at terminal time \(T\)
\[
\Bar{\text{AVaR}}_t^\alpha (L_t^2) = \frac{1}{1 - \alpha} \int_0^1 F_Z^{-1}(u) \, d\mathcal{P}_t \left( \frac{a_1}{1 - \alpha} \int_0^1 F_Z^{-1}(u)^2 \, du + b \right), \quad t = 0, \ldots, T - 1,
\]
where \(\mathcal{P}_t : \mathbb{R} \to \mathbb{R}\) is an \(\mathcal{F}_t\)-measurable mapping given by
\[
\mathcal{P}_t(x) = a_0 \sum_{k=0}^{T-t-2} x^k + \sigma_{t+1}^2 x^{T-t-1}, \quad t = 0, \ldots, T - 1.
\]

4.2 Almost sure bounds for AVaR

4.2.1 AVaR-bounds for single day losses

Finding an analytical expression for \(\Bar{\text{AVaR}}_t^\alpha (L)\) for the (unsquared) GARCH(1,1) loss is not straightforward. It is however possible to derive a closed form approximation for \(\Bar{\text{AVaR}}_t^\alpha (L_T)\). To this end, we need the following lemma.

Lemma 4.6. For \(t = 0, \ldots, T - 2\)
\[
\{\omega \in \Omega : \sigma_{t+2} > \text{VaR}_t^\alpha (\sigma_{t+2})\} = \{\omega \in \Omega : Z_{t+1} > F_Z^{-1}(\alpha)\}.
\]

Proof. From \((3.3)\) we know that
\[
\text{VaR}_t^\alpha (\sigma_{t+2}) = \sqrt{a_0 + \sigma_{t+1}^2 (a_1 F_Z^{-1}(\alpha))^2 + b}.
\]

From the definition of the GARCH(1,1) model \((2.5)\) we conclude
\[
\sigma_{t+2}^2 = a_0 + \sigma_{t+1}^2 (a_1 Z_{t+1}^2 + b) > a_0 + \sigma_{t+1}^2 (a_1 F_Z^{-1}(\alpha))^2 + b = \text{VaR}_t^\alpha (\sigma_{t+2})^2
\]
if and only if \(Z_{t+1} > F_Z^{-1}(\alpha)\). \(\square\)

We now derive closed form bounds for \(\Bar{\text{AVaR}}_t^\alpha\). We begin with the upper bound \(\Bar{\text{AVaR}}_t^\alpha (L_T)\) which arises from an application of Jensen’s inequality. For the proofs of the following two Propositions we refer to Appendix B.

Proposition 4.7. Let \(L_T\) be the loss position at time \(T > 0\) given by the GARCH(1,1) model \((2.5)\). Let \(\alpha \in (0,1)\) be such that \(F_Z^{-1}(\alpha) > 0\). Then
\[
\Bar{\text{AVaR}}_t^\alpha (L_T) := \frac{1}{1 - \alpha} \int_0^1 F_Z^{-1}(y) \, dy \sqrt{\mathcal{P}_t \left( \frac{a_1}{1 - \alpha} \int_0^1 F_Z^{-1}(u)^2 \, du + b \right)}, \quad t = 0, \ldots, T - 1,
\]
where \(\mathcal{P}_t(\cdot)\) is given by \((3.8)\), satisfies
\[
\text{AVaR}_t^\alpha \leq \Bar{\text{AVaR}}_t^\alpha \quad t = 0, \ldots, T - 1.
\]
In Section 5.3 we investigate the accuracy of the upper bound \((4.5)\). An easy alteration of the proof of the previous result yields a closed form lower bound \(\text{AVaR}_t^\alpha\) for \(\Bar{\text{AVaR}}_t^\alpha\).

Proposition 4.8. Let \(L_T\) be the loss position at time \(T > 0\) given by the GARCH(1,1) model \((2.5)\). Let \(\alpha \in (0,1)\) be such that \(F_Z^{-1}(\alpha) > 0\). Then
\[
\text{AVaR}_t^\alpha (L_T) := \frac{1}{1 - \alpha} \int_0^1 F_Z^{-1}(u) \, du \left( \frac{1}{1 - \alpha} \int_0^1 \sqrt{a_1 F_Z^{-1}(y)^2 + b} \, dy \right)^{T-t-1} \sigma_{t+1}, \quad t = 0, \ldots T - 1,
\]
satisfies
\[
\text{AVaR}_t^\alpha (L_T) \geq \Bar{\text{AVaR}}_t^\alpha (L_T) \quad t = 0, \ldots, T - 1.
\]
Similar to Corollary 3.8 we can easily compute \( \overline{\text{AVaR}}^\alpha \) and \( \overline{\text{AVaR}}^\alpha \) by scaling the 1-day-ahead AVaR\(^\alpha\). This is important as it allows a deterministic and fast computation of lower and upper AVaR bounds. The following Corollary is straightforward.

**Corollary 4.9.** Let the assumptions from Propositions 4.7 and 4.8 be in force. Then for fixed \( t \in \{1, \ldots, T - 1\} \) and \( m \in \mathbb{N} \) such that \( t + m \leq T \) we have

\[
\overline{\text{AVaR}}_t^{\alpha}(L_{t+m+1}) = \frac{a_{t+1}}{1-\alpha} \int_\alpha^1 F_Z^{-1}(y)dy \left( \frac{a_1}{1-\alpha} \int_\alpha^1 F_Z^{-1}(u)^2du + b \right)^{m/2},
\]

and

\[
\overline{\text{AVaR}}_t^{\alpha}(L_{t+m+1}) = \frac{a_{t+1}}{1-\alpha} \int_\alpha^1 F_Z^{-1}(y)dy \left( \frac{1}{1-\alpha} \int_\alpha^1 \sqrt{a_1 F_Z^{-1}(y)^2 + b}dy \right)^m.
\]

**Remark 4.10.** Using (4.7) and (4.8), we have a hindsight on the gap between the lower and upper bound. Given the loss position \( L_{t+m+1} \) we can check whether the quantity

\[
\left( \frac{1}{1-\alpha} \int_\alpha^1 \sqrt{a_1 F_Z^{-1}(y)^2 + b}dy \right)^m
\]

is close to 1.

In Section 5.3 we study the accuracy of the lower bound (4.6). We also study the tighteness of \( \overline{\text{AVaR}}^\alpha \) and \( \overline{\text{AVaR}}^\alpha \) by checking the gap from Remark 4.10.

### 4.2.2 AVaR-bounds for aggregated losses

Again we can compute the \( m \)-day ahead AVaR\(^\alpha\).

**Proposition 4.11.** Let the assumptions of Theorem 4.5 hold, then for \( t \in \{1, \ldots, T - 1\} \) and \( m \in \mathbb{N} \) such that \( t + m \leq T \) the \( m \)-day-ahead \( \overline{\text{AVaR}}_t^{\alpha} \) of aggregated losses \( \sum_{k=1}^m L_{t+k} \) is bounded by

\[
\sum_{k=1}^m \overline{\text{AVaR}}_t^{\alpha}(L_{t+m+1}) \leq \overline{\text{AVaR}}_t^{\alpha}(\sum_{k=1}^m L_{t+m+1}) \leq \sum_{k=1}^m \overline{\text{AVaR}}_t^{\alpha}(L_{t+m+1}).
\]

**Proof.** To prove (4.9), note that according to Corollary 2.5 \( \overline{\text{AVaR}}_t^{\alpha}(\sum_{k=1}^m L_{t+m+1}) \) satisfies

\[
\overline{\text{AVaR}}_t^{\alpha}(\sum_{k=1}^m L_{t+m+1}) = \sum_{k=1}^m \overline{\text{AVaR}}_t^{\alpha}(L_{t+m+1}).
\]

Now the assertion follows from an application of Propositions 4.7 and 4.8.

### 5 Extreme value theory based quantile estimation

#### 5.1 Generalized Pareto Distribution

Up to now we have not fixed the noise distribution, only assumed sometimes that it has a density or a continuous distribution function. Throughout we worked with \( \alpha \) close to 1 corresponding to the noise distribution function to be close to 1. Thus it is sufficient to specify the distribution function above some high threshold \( u \). This is a typical assumption in extreme value theory, and we will apply the Peaks-over-Threshold method (as in McNeil and Frey (2000)). We first explain the setting. The *Generalized Pareto distribution (GPD)* is given by

\[
G_{\xi,\beta}(x) = \begin{cases} 
1 - \left(1 + \frac{x}{\beta}\right)^{-1/\xi}, & \xi \neq 0, \\
1 - \exp\left( -\frac{x}{\beta}\right), & \xi = 0,
\end{cases}
\]

(5.1)
where $\beta > 0$ and $\xi \in \mathbb{R}$. If $\xi > 0$ (5.1) is defined for $x \geq 0$ and if $\xi < 0$ (5.1) is defined on $x \in [0, -\beta/\xi]$, see e.g. Section 3.4 in Embrechts et al. (1997). Assume that we fix some high threshold $u > 0$. Given a random variable $X$ with distribution function $F$ and right endpoint $x_F$, its associated excess distribution function is defined as

$$F_u(y) = \mathbb{P}(X - u \leq y \mid X > u) = \frac{F(y + u) - F(u)}{1 - F(u)}, \quad 0 \leq y < x_F - u.$$  

The strength of the GPD is compressed in a result by Pickands (1975) and Balkema and de Haan (1974) which classifies the GPD as the limit distribution of a large class of excess distributions. More precisely, under mild conditions there exists a measurable non-negative parameter $\beta = \beta(u)$ such that

$$\lim_{u \to x_F} \sup_{0 \leq y \leq x_F - u} |F_u(y) - G_{\xi,\beta}(u)| = 0$$

holds, see Theorem 3.4.13 in Embrechts et al. (1997) for a rigorous statement of this result. The density of (5.1) is given by

$$g_{\xi,\beta}(x) = \begin{cases} \frac{1}{\beta}(1 + \xi x/\beta)^{-1/\xi - 1}, & \xi \neq 0, \\ \frac{1}{\beta} \exp\left(-x/\beta\right), & \xi = 0. \end{cases}$$

Under the assumption that $Z$ has the distribution function $F_Z$, which for some high enough threshold $u > 0$ satisfies $F_u(x) = G_{\xi,\beta}(x)$ for $0 \leq x \leq x_F - u$ and for some $\xi \in \mathbb{R}$ and $\beta > 0$, we find for $\alpha \geq F(u)$

$$F_Z^{-1}(\alpha) = u + \frac{\beta}{\xi}\left(\left(1 - \frac{\alpha}{1 - F(u)}\right)^{\xi} - 1\right),$$

and

$$\frac{1}{1 - \alpha} \int_0^1 F_Z^{-1}(y)dy = \frac{F_Z^{-1}(\alpha)}{1 - \xi} + \frac{\beta - \xi u}{1 - \xi}.$$

Figure 1: Motorola stock analysis: loss data (top), conditional variances after fitting a GARCH(1,1) model (middle), and residuals of loss data (bottom).
Note also that the antiderivative of $F_Z^{-1}(x)^2$ is given by

$$
\frac{d}{dx} F_Z^{-1}(x)^2 = (\beta/\xi - u)^2 x + \frac{(\beta/\xi)^2 (\frac{1-x}{1-F(u)})^{-2\xi}}{1-2\xi} + 2 \frac{\beta(u-\beta/\xi)(\frac{1-x}{1-F(u)})^{-\xi}}{\xi(1-\xi)}
$$

which readily allows us to find an analytical expression for $\int_\alpha^1 F_Z^{-1}(x)^2dx$ by integration by parts.

### 5.2 Statistical model fitting

In this section we apply the theory and formulas derived previously to a data set. We choose the historical daily closing prices of the Motorola stock from 1st March 1985 until 15th October 2014 as this data set provides several canonical features of financial time series. We transform prices into negative log-returns; i.e., into losses, and fit the GARCH(1,1) parameters using Quasi Maximum Likelihood Estimation (QMLE) (e.g. [Franq and Zakoian (2010)], Chapter 7). The parameter estimates can be found in Table 1, and the outcome is depicted in Figure 1.

We see in the middle plot major clustering of volatility in October 1987 (Black Monday), in a pronounced period between 2000 until 2002 (Dot-com bubble and wake of 9/11 attacks) and in a longer lasting period following the financial crisis between 2008 until 2010.

In a next step we examine the sample autocorrelation functions of the loss data and the residuals after fitting a GARCH(1,1) model. In Figure 2 the bottom plots depict the acf of the residuals and the squared residuals and is supportive for the our assumption of i.i.d. GARCH residuals $Z_t$. This is also reflected in several runs of the Ljung-Box for various lags for the residuals. The residuals also pass the augmented Dickey-Fuller and the KPSS stationarity tests.

As explained in Section 5.1 fit a GPD to the upper tail of the residuals. We first have to choose a high enough threshold value $u$ and we choose it as the approximate 92% quantile of the residuals.

Figure 2: Motorola stock analysis: sample autocorrelation functions for loss data (top) and residuals after fitting a GARCH(1,1) model (bottom).
Figure 3: Fit of the Generalized Pareto Distribution. Left: mean excess plot of the positive residuals, and QQ-plot of the threshold exceeding residuals against the fitted GPD. Right: excess distribution $F_u(x-u)$ from the fitted GPD model (solid line) against the empirical estimates of excess probabilities (dotted points).

Figure 4: QQ-plot of the threshold exceeding residuals against the fitted GPD.

This is supported by studying the mean excess plot of the nonnegative residuals in Figure 3. The 92% quantile of the residuals (solid blue line) yields a threshold which sufficiently marks the beginning of the linear behaviour of the mean excess plot. Since the empirical mean excesses are increasing, we may assume that the shape parameter $\xi$ is positive. This is confirmed by the parameter estimates for $\xi$ and $\beta$. The Maximum Likelihood Estimators are $\hat{\xi} = 0.3376$ with a 95% confidence interval $[0.2272, 0.4481]$ and $\hat{\beta} = 0.4609$ with a 95% confidence interval $[0.4023, 0.5280]$.

In Figure 3, the right hand plot depicts the GPD fit of the excess distribution $F_u(x-u) = \mathbb{P}(X \leq x \mid X > u)$ superimposed on empirical estimates of excess probabilities. Note how well the GPD model fits to the empirical estimates of the excess probabilities.

| Parameter | Value | Standard error |
|-----------|-------|----------------|
| $a_0$     | 2e-07 | 1.09e-07       |
| $\hat{a}_1$ | 0.0451 | 0.0014        |
| $\hat{\beta}$ | 0.9531 | 0.0013        |

Table 1: Estimated GARCH(1,1) parameters by QMLE.
A QQ-plot of the empirical quantiles against the fitted quantiles is depicted in Figure 4. Note again the good correspondence of the fitted GPD with the empirical estimates.

5.3 Testing time-consistent risk measures on data

We now compute the corresponding time-consistent risk measures based on Sections 3 and 4. In a first step, we compute the \( m \)-day-ahead time-consistent VaR given by Proposition 3.7 for different levels of \( \alpha \); i.e., we fix \( t \) and consider \( \overline{\text{VaR}}^\alpha_t(L_{t+m}) \) for various \( m \in \mathbb{N} \). In Figure 5 we plot \( \overline{\text{VaR}}^\alpha_t(L_{t+m}) \) for \( m = 1, \ldots, 10 \). Once the single time consistent risk measures \( \overline{\text{VaR}}^\alpha_t(L_{t+m}) \) are computed, we simultaneously get the risk measure of the aggregated losses over \( m \) days from Proposition 3.9 by aggregation; i.e.,

\[
\overline{\text{VaR}}^\alpha_t \left( \sum_{j=1}^{m} L_{t+j} \right) = \sum_{j=1}^{m} \overline{\text{VaR}}^\alpha_j(L_{t+j}).
\]

Analogously we compute the approximate AVaR bounds for single and for aggregated losses according to Propositions 4.7 and 4.8. In Figure 7 we see that for large values of \( m \) and \( \alpha \) close to 1 the discrepancy between \( \overline{\text{AVaR}}^\alpha \) and \( \overline{\text{AVaR}}^{\alpha} \) becomes increasingly larger.

This is emphasized by Figure 8 which shows in the first two plots the risk measure surface for single losses \( L_{t+m} \) for different values for \( \alpha \) and \( m \), and the third plot depicts the difference \( \overline{\text{AVaR}}^\alpha - \overline{\text{AVaR}}^{\alpha} \); we see that the larger \( \alpha \) and \( m \) the greater the difference \( \overline{\text{AVaR}}^\alpha - \overline{\text{AVaR}}^{\alpha} \) becomes. This shows that the upper bound \( \overline{\text{AVaR}}^{\alpha} \) does not accurately capture the time-consistent evaluation of \( \overline{\text{AVaR}}^\alpha \).

As before we can now compute lower and upper bounds for the time-consistent AVaR for aggregated losses \( \sum_{j=1}^{m} L_{t+j} \) by means of Proposition 2.5 and Corollary 4.9. The outcome is depicted in Figure 9. As expected, aggregating the single loss positions over time yields a similar qualitative behavior as we can see in Figure 8.

Table 2 shows the values of \( \overline{\text{VaR}}^\alpha_t \) for single losses \( L_{t+m} \) and aggregated losses \( \sum_{j=1}^{m} L_{t+j} \) for various levels of \( \alpha \) and \( m = 1, \ldots, 10 \). Note that the risk measure for the aggregated losses arise from the cumulative summation of the risk measures for single losses.

In Table 3 we see a comparison between the lower bounds \( \overline{\text{AVaR}}^\alpha \) and the upper bounds \( \overline{\text{AVaR}}^{\alpha} \). As already noted the higher the values for \( m \) and \( \alpha \) the greater the difference between the lower

![Figure 5: Time-consistent VaR estimation for single loss positions \( m \) days ahead.](image)
Figure 6: Time-consistent VaR estimation for aggregated loss positions $m$ days ahead.

Figure 7: $\overline{AVaR}_\alpha$ and $\underline{AVaR}_\alpha$ for $\alpha = 97.5\%$ and $\alpha = 99\%$ for single loss positions.

Figure 8: $\overline{AVaR}_\alpha$, $\underline{AVaR}_\alpha$ and $\overline{AVaR}_\alpha - \underline{AVaR}_\alpha$ for single loss positions.
and the upper bound becomes. A comparison the single losses $\overline{\text{VaR}}^{\alpha}_t(L_{t+m})$ with $\text{AVaR}^{\alpha}_t(L_{t+m})$ however yields that $\overline{\text{VaR}}^{\alpha}_t(L_{t+m}) \leq \text{AVaR}^{\alpha}_t(L_{t+m})$ holds in this experiment; i.e., the lower bound of the time-consistent AVaR always dominates the time-consistent VaR.

The discrepancy between lower and upper bounds becomes more pronounced when we consider risk measures for aggregated losses. In Table 4 the values for $\text{AVaR}^{\alpha}$ and $\overline{\text{AVaR}}^{\alpha}$ arise by summing
cumulatively the risk values of the single day losses from Table 3. This summation enforces the gap between \( \text{AVaR}_t^\alpha (\sum_{j=1}^m L_{t+j}) \) and \( \tilde{\text{AVaR}}_t^\alpha (\sum_{j=1}^m L_{t+j}) \) as \( m \) becomes large: since the discrepancy between the single loss positions \( \text{AVaR}_t^\alpha (L_{t+m}) \) and \( \tilde{\text{AVaR}}_t^\alpha (L_{t+m}) \) are already large, the summation of all previous risk measures hence amplifies the gap.

### A Proofs of Section 3

**Proof of Theorem 3.7** We proceed by backward induction. Denote \( \kappa := F_T^{-1}(\alpha) \) then, by (3.2), at \( T-1 \) we have the 1-day-ahead-VaR

\[
\text{VaR}^\alpha_{T-1}(L_T) = \kappa \sigma_T
\]

which agrees with (3.7) for \( t = T-1 \). Assume that (3.7) holds for all \( s = t, \ldots, T-1 \). We have

\[
\text{VaR}^\alpha_{T-1}(L_T) = \text{VaR}^\alpha_{T-1}(\tilde{\text{VaR}}^\alpha_{T-1}(L_T)) = \text{VaR}^\alpha_{T-1}(\kappa \sqrt{P_t(a_1 \kappa^2 + b)}) = \text{essinf}\{m \in L^0(\mathcal{F}_{t-1}) : P(\kappa \sqrt{P_t(a_1 \kappa^2 + b)} \leq m | \mathcal{F}_{t-1}) \geq \alpha\}.
\]

Note that

\[
P(\kappa \sqrt{P_t(a_1 \kappa^2 + b)} \leq m | \mathcal{F}_{t-1}) = P(\kappa^2 P_t(a_1 \kappa^2 + b) \leq m^2 | \mathcal{F}_{t-1}) = P\left(\sigma^2_{t+1}(a_1 \kappa^2 + b)^{T-t-1} \leq \left(\frac{m^2}{\kappa}\right)^2 - a_0 \sum_{k=0}^{T-t-2} (a_1 \kappa^2 + b)^k | \mathcal{F}_{t-1}\right).
\]

Using the definition of the GARCH volatility (2.5) for \( \sigma^2_{t+1} \) this can be continued by

\[
P(\kappa \sqrt{P_t(a_1 \kappa^2 + b)} \leq m | \mathcal{F}_{t-1}) = P(a_1 \sigma^2_t Z_t^2 \leq \frac{1}{(a_1 \kappa^2 + b)^{T-t-1}} \left(\left(\frac{m}{\kappa}\right)^2 - a_0 \sum_{k=0}^{T-t-2} (a_1 \kappa^2 + b)^k\right) - a_0 - b \sigma^2_t | \mathcal{F}_{t-1}) = P\left(Z_t \leq \frac{1}{a_1 \sigma_t} \left\{ \frac{1}{(a_1 \kappa^2 + b)^{T-t-1}} \left(\left(\frac{m}{\kappa}\right)^2 - a_0 \sum_{k=0}^{T-t-2} (a_1 \kappa^2 + b)^k\right) - a_0 - b \sigma^2_t \right\}^{1/2} | \mathcal{F}_{t-1}\right).
\]
Since \( \sigma_t \) is \( \mathcal{F}_{t-1} \)-measurable and \( Z_t \) is independent of \( \mathcal{F}_{t-1} \) we conclude that

\[
\tilde{\text{VaR}}_{t-1}^\alpha(L_T) = \kappa \sqrt{\left(a_0 + (\kappa^2 a_1 + b)\sigma_t^2 \right)(a_1 \kappa^2 + b) T^{-1} + a_0 \sum_{k=0}^{t-2} (a_1 \kappa^2 + b)^k}
\]

\[
= \kappa \sqrt{a_0 \sum_{k=0}^{T-t-1} (a_1 \kappa^2 + b)^k + \sigma_t^2 (a_1 \kappa^2 + b) T^{-t}}
\]

\[
= \kappa \sqrt{\mathcal{P}_{t-1}(a_1 \kappa^2 + b)}.
\]

This finishes the proof. \( \square \)

### B Proofs of Section 4

**Proof of Theorem 4.5** We apply again backward induction. Since, by (3.2), \( \text{VaR}_{T-1}^\alpha(Z_T^2) = F^{-1}_Z(\alpha)^2 \), at \( T-1 \) we have

\[
\tilde{\text{VaR}}_{T-1}^\alpha(L_T) = \sigma_T^2 \mathbb{E}[Z_T^2 | Z_T^2 > \text{VaR}_{T-1}^\alpha(Z_T^2), \mathcal{F}_{T-1}].
\]

Hence,

\[
\mathbb{E}[Z_T^2 | Z_T^2 > \text{VaR}_{T-1}^\alpha(Z_T^2)] = \frac{1}{1-\alpha} \int_{F^{-1}_Z(\alpha)}^\infty u^2 dF_Z(u) = \frac{1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du.
\]

Thus,

\[
\tilde{\text{VaR}}_{T-1}^\alpha(L_T) = \frac{\sigma_T^2}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du,
\]

which agrees with (4.4) for \( t = T-1 \). Let us assume that (4.4) holds for all \( s = t, \ldots, T-1 \). Then it remains to prove (4.4) for \( t-1 \). We have

\[
\tilde{\text{VaR}}_{t-1}^\alpha(L_T^2) = \text{VaR}_{t-1}^\alpha(\tilde{\text{VaR}}_{t-1}^\alpha(L_T^2))
\]

\[
= \text{VaR}_{t-1}^\alpha \left( \frac{1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du \right) \mathcal{P}(\frac{a_1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du + b).
\]

We denote \( G_t := \mathcal{P}(\frac{a_1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du + b) \) and take out the constant, which yields

\[
\tilde{\text{VaR}}_{t-1}^\alpha(L_T^2) = \frac{1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du \mathbb{E}[G_t | G_t > \text{VaR}_{t-1}^\alpha(G_t), \mathcal{F}_{t-1}].
\]

Now note that by Definition 3.1

\[
\text{VaR}_{t-1}^\alpha(G_t) = \text{essinf}\{m \in L^0(\mathcal{F}_{t-1}) : \mathbb{P}(G_t \leq m | \mathcal{F}_{t-1}) > \alpha\}.
\]

We compute further, using the definition of the GARCH volatility (2.5) for \( \sigma_{t+1}^2 \)

\[
\mathbb{P}(G_t \leq m | \mathcal{F}_{t-1}) = \mathbb{P}(a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du + b \right)^k + \sigma_{t+1}^2 \left( \frac{a_1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du + b \right)^{T-t-1} \leq m | \mathcal{F}_{t-1})
\]

\[
= \mathbb{P}(\sigma_{t+1} \leq \frac{m - a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du + b \right)^k}{\left( \frac{a_1}{1-\alpha} \int_\alpha^1 F^{-1}_Z(u)^2 du + b \right)^{T-t-1}} | \mathcal{F}_{t-1})
\]

19
\[ \mathbb{P}(Z_t^2 \leq \left( \frac{m - a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^k}{a_1 \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b} \right)^{T-t-1} - a_0 - b\sigma_t^2 \alpha_1 \sigma_t^2), \]

where in the last line we have used that \( \sigma_t \) is \( \mathcal{F}_{t-1} \)-measurable and the independence of \( Z_t \) and \( \mathcal{F}_{t-1} \).

We can thus conclude that

\[
\text{VaR}_{t-1}^\alpha(G_t) = \left( a_0 + a_1 \sigma_t^2 F_Z^{-1}(\alpha)^2 + b\sigma_t^2 \right) \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^{T-t-1} + a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^k.
\]

From Lemma 4.6 we know that \( \{G_t > \text{VaR}_{t-1}^\alpha(G_t)\} = \{Z_t > F_Z^{-1}(\alpha)\} \). Hence, it follows from the independence of \( Z_t \) and \( \mathcal{F}_{t-1} \) that

\[
\mathbb{E}[G_t \mid G_t > \text{VaR}_{t-1}^\alpha(G_t), \mathcal{F}_{t-1}] = \mathbb{E}[G_t \mid Z_t > F_Z^{-1}(\alpha)] = \frac{1}{1-\alpha} \mathbb{E}[G_t \mathbb{1}_{\{Z_t > F_Z^{-1}(\alpha)\}}]. \tag{B.1}
\]

Moreover, we calculate

\[
\mathbb{E}\left[G_t \mathbb{1}_{\{Z_t > F_Z^{-1}(\alpha)\}}\right] = (1-\alpha) a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^k + \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^{T-t-1} \int_{\alpha}^\infty \left( a_0 + a_1 \sigma_t^2 u^2 + b\sigma_t^2 \right) dF_Z(u)
\]

\[
= (1-\alpha) a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^k + \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^{T-t-1} \int_{\alpha}^1 \left( a_0 + a_1 \sigma_t^2 F_Z^{-1}(u)^2 + b\sigma_t^2 \right) du
\]

\[
= (1-\alpha) a_0 \sum_{k=0}^{T-t-2} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^k + \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^{T-t-1} \left( a_1 \int_{\alpha}^1 F_Z^{-1}(u)^2 du + (1-\alpha)b \right) \sigma_t^2,
\]

which in combination with (B.1) yields

\[
\mathbb{E}[G_t \mid G_t > \text{VaR}_{t-1}^\alpha(G_t), \mathcal{F}_{t-1}] = a_0 \sum_{k=0}^{T-t-1} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^k + \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right)^{T-t} \sigma_t^2
\]

\[
= \mathcal{P}_{t-1} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right).
\]

This finally amounts to

\[
\text{AVaR}_{t-1}^\alpha(L^2_t) = \frac{1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u) du \mathbb{E}[G_t \mid G_t > \text{VaR}_{t-1}^\alpha(G_t), \mathcal{F}_{t-1}] \]

\[
= \frac{1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u) du \mathcal{P}_{t-1} \left( \frac{a_1}{1-\alpha} \int_{\alpha}^1 F_Z^{-1}(u)^2 du + b \right),
\]

which proves the assertion. \(\square\)
Proof of Proposition 4.7. A careful proof tracking reveals its similarity to the proof of Theorem 4.5. At \( t = T - 1 \) we have \( \overline{\text{AVaR}}_{T-1}^\alpha(L_T) = \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y) dy \sigma_T \) which coincides with \( \text{AVaR}_{T-1}^\alpha(L_T) \). Since

\[ \overline{\text{AVaR}}_{T-2}^\alpha(L_T) = \text{AVaR}_{T-2}^\alpha(\overline{\text{AVaR}}_{T-1}^\alpha(L_T)) = \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y) dy \text{AVaR}_{T-2}^\alpha(\sigma_T), \]

and since the definition of \( \text{AVaR} \) yields

\[ \text{AVaR}_{T-2}^\alpha(\sigma_T) = \mathbb{E} \left[ \sigma_T \mid \sigma_T > \text{VaR}_{T-2}^\alpha(\sigma_T), \mathcal{F}_{t-2} \right] = \mathbb{E} \left[ \sigma_T \mid Z_{T-1} > F_Z^{-1}(\alpha), \mathcal{F}_{t-2} \right], \]

by Lemma 4.6 an application of Jensen’s inequality yields

\[ \mathbb{E} \left[ \sigma_T \mid Z_{T-1} > F_Z^{-1}(\alpha), \mathcal{F}_{t-2} \right] \leq \left( \mathbb{E} \left[ \sigma_T^2 \mid Z_{T-1} > F_Z^{-1}(\alpha), \mathcal{F}_{t-2} \right] \right)^{1/2}. \]

We obtain further

\[ \mathbb{E} \left[ \sigma_T^2 \mid Z_{T-1} > F_Z^{-1}(\alpha), \mathcal{F}_{t-2} \right] = \mathbb{E} \left[ a_0 + a_1 \sigma_{T-1}^2 Z_{T-1}^2 + b \sigma_{T-1}^2 \mid Z_{T-1} > F_Z^{-1}(\alpha), \mathcal{F}_{t-2} \right] \]

\[ = \frac{1}{1-\alpha} \int_{F_Z^{-1}(\alpha)}^\infty \left( a_0 + a_1 \sigma_{T-1}^2 (a_1 y^2 + b) \right) dF_Z(y) = a_0 + b \sigma_{T-1}^2 + \frac{a_1}{1-\alpha} \sigma_{T-1}^2 \int_{F_Z^{-1}(\alpha)}^\infty y^2 dF_Z(y) \]

\[ = a_0 + b \sigma_{T-1}^2 + \frac{a_1}{1-\alpha} \sigma_{T-1}^2 \int_0^1 F_Z^{-1}(y)^2 dy = a_0 + a_1 \int_0^1 F_Z^{-1}(y)^2 dy + b \]

which amounts to

\[ \overline{\text{AVaR}}_{T-2}^\alpha(L_T) \leq \left( \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y)^2 dy \sigma_T^2 \right)^{1/2} \]

\[ = \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y) dy \sigma_T^2 \left( \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y)^2 dy + b \right) \]

\[ = \text{AVaR}_{T-1}^\alpha(L_T). \]

This proves for \( t = T - 2 \) that \( \overline{\text{AVaR}}_{T-1}^\alpha(L_T) \) is an upper bound for \( \overline{\text{AVaR}}_{T-2}^\alpha(L_T) \).

Now assume that \( \overline{\text{AVaR}}_{s}^\alpha(L_T) \geq \overline{\text{AVaR}}_{s}^\alpha(L_T) \) holds true for \( s = T - 1, \ldots, t + 1 \). We show next that also

\[ \overline{\text{AVaR}}_{s}^\alpha(L_T) \geq \overline{\text{AVaR}}_{s}^\alpha(L_T). \]

To this end notice that

\[ \overline{\text{AVaR}}_{s}^\alpha(L_T) = \text{AVaR}_{s}^\alpha(\overline{\text{AVaR}}_{s+1}^\alpha(L_T)) \leq \text{AVaR}_{s}^\alpha(\overline{\text{AVaR}}_{s+1}^\alpha(L_T)). \]

Moreover, we have

\[ \text{AVaR}_{s}^\alpha(\overline{\text{AVaR}}_{s+1}^\alpha(L_T)) = \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y) dy \text{AVaR}_{s}^\alpha \left( \sqrt{P_{s+1} \left( \frac{a_1}{1-\alpha} \int_0^1 F_Z^{-1}(u)^2 du + b \right)} \right) \]

\[ = \frac{1}{1-\alpha} \int_0^1 F_Z^{-1}(y) dy \mathbb{E} \left[ \sqrt{P_{s+1} \left( \frac{a_1}{1-\alpha} \int_0^1 F_Z^{-1}(u)^2 du + b \right)} \mid \mathcal{F}_{t-1} \right], \]

\[ \sqrt{P_{s+1} \left( \frac{a_1}{1-\alpha} \int_0^1 F_Z^{-1}(u)^2 du + b \right)} > \text{VaR}_{s}^\alpha \left( \sqrt{P_{s+1} \left( \frac{a_1}{1-\alpha} \int_0^1 F_Z^{-1}(u)^2 du + b \right)} \right). \]
By a similar calculation as in the proof of Theorem 4.5, we can see that the above expression simplifies to

\[
\text{AVaR}_\alpha^o(\text{AVaR}_{t+1}^o(L_T)) = \frac{1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \left( \text{P}_{t+1} \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \right) \left( 1 - \text{P}_{t+1} \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \right)^{1/2},
\]

where the last line follows from Jensen’s inequality. Note that

\[
\text{AVaR}_\alpha(\text{AVaR}_{t+1}^o(L_T)) = \frac{1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \left( \text{P}_{t+1} \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \right) \left( 1 - \text{P}_{t+1} \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \right)^{1/2},
\]

which implies

\[
\text{AVaR}_\alpha^o(\text{AVaR}_{t+1}^o(L_T)) \leq \frac{1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(y) dy \sqrt{\text{P}_{t+1} \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \left( 1 - \text{P}_{t+1} \left( \frac{a_1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right) \right)} = \text{AVaR}_\alpha^o(L_T).
\]

Finally it follows from (B.2) that \( \text{AVaR}_\alpha^o(L_T) \leq \text{AVaR}_\alpha^o(L_T) \). \( \square \)

**Proof of Proposition 4.8**  
At \( t = T - 1 \), \( \text{AVaR}_{T-1}^o(L_T) \) coincides with \( \text{AVaR}_{T-1}^o(L_T) \). At \( t = T - 2 \) note that

\[
\text{AVaR}_{T-2}^o(L_T) = \text{AVaR}_{T-2}^o(\text{AVaR}_{T-1}^o(L_T)) = \frac{1}{1 - \alpha} \int_\alpha^1 F_{Z_t}^{-1}(y) dy \text{AVaR}_{T-2}^o(\sigma_T).
\]

By the definition of AVaR

\[
\text{AVaR}_{T-2}^o(\sigma_T) = \mathbb{E} \left[ \sigma_T \mid \sigma_T > \text{VaR}_{T-2}^o(\sigma_T), \mathcal{F}_{t-2} \right],
\]

which by Lemma 4.6 rewrites as

\[
\text{AVaR}_{T-2}^o(\sigma_T) = \mathbb{E} \left[ \sigma_T \mid Z_{T-1} > F_{Z_t}^{-1}(\alpha), \mathcal{F}_{t-2} \right].
\]

We have

\[
\mathbb{E} \left[ \sigma_T \mid Z_{T-1} > F_{Z_t}^{-1}(\alpha), \mathcal{F}_{t-2} \right] = \mathbb{E} \left[ \sqrt{a_0 + \sigma_T^2 \left( a_1 \int_\alpha^1 F_{Z_t}^{-1}(u) du + b \right)} \mid Z_{T-1} > F_{Z_t}^{-1}(\alpha), \mathcal{F}_{t-2} \right]
\]

22
\[
\begin{align*}
&\geq \sigma_{T-1} \mathbb{E}\left[ \sqrt{a_1 Z_{T-1}^2 + b} \mid Z_{T-1} > F_Z^{-1}(\alpha), \mathcal{F}_{t-2} \right] \\
&= \sigma_{T-1} \frac{1}{1 - \alpha} \int_{F_Z^{-1}(\alpha)}^{\infty} \sqrt{a_1 y^2 + b} dF_Z(y),
\end{align*}
\]
where the inequality follows from the fact that \( a_0 > 0 \). Hence it follows that
\[
\text{AVaR}_{T-2}(L_T) \geq \frac{1}{1 - \alpha} \int_{\alpha}^{1} F_Z^{-1}(y) dy \frac{1}{1 - \alpha} \int_{F_Z^{-1}(\alpha)}^{\infty} \sqrt{a_1 y^2 + b} dF_Z(y) \sigma_{T-1} = \text{AVaR}_{T-2}(L_T).
\]
Now assume that \( \text{AVaR}^\alpha_t(L_T) \leq \text{AVaR}^\alpha_{t+1}(L_T) \) holds true for \( s = T - 1, \ldots, t + 1 \). Let us show that we also have
\[
\text{AVaR}^\alpha_t(L_T) \leq \text{AVaR}^\alpha_{t+1}(L_T).
\]
To this end notice that
\[
\text{AVaR}^\alpha_t(L_T) = \text{AVaR}^\alpha_t(\text{AVaR}^\alpha_{t+1}(L_T)) \geq \text{AVaR}^\alpha_t(\text{AVaR}^\alpha_{t+1}(L_T)).
\]
We have
\[
\text{AVaR}^\alpha_t(\text{AVaR}^\alpha_{t+1}(L_T)) = \frac{1}{1 - \alpha} \int_{\alpha}^{1} F_Z^{-1}(u) du \left( \frac{1}{1 - \alpha} \int_{F_Z^{-1}(\alpha)}^{\infty} \sqrt{a_1 y^2 + b} dF_Z(y) \right)^{T-t-2} \mathbb{E}[\sigma_{t+1} \mid \sigma_{t+2} > \text{VaR}^\alpha_t(\sigma_{t+2}), \mathcal{F}_t] \\
= \frac{1}{1 - \alpha} \int_{\alpha}^{1} F_Z^{-1}(u) du \left( \frac{1}{1 - \alpha} \int_{F_Z^{-1}(\alpha)}^{\infty} \sqrt{a_1 y^2 + b} dF_Z(y) \right)^{T-t-2} \mathbb{E}[\sigma_{t+1} \mid Z_{t+1} > F_Z^{-1}(\alpha), \mathcal{F}_t],
\]
where the last equality follows from Lemma 1.6. Since
\[
\mathbb{E}[\sigma_{t+2} \mid Z_{t+1} > F_Z^{-1}(\alpha), \mathcal{F}_t] = \mathbb{E}\left[ \sqrt{a_0 + \sigma_{t+1}^2(a_1 Z_{t+1}^2 + b)} \mid Z_{t+1} > F_Z^{-1}(\alpha), \mathcal{F}_t \right] \\
\geq \sigma_{t+1} \mathbb{E}\left[ \sqrt{a_1 Z_{t+1}^2 + b} \mid Z_{t+1} > F_Z^{-1}(\alpha), \mathcal{F}_t \right] \\
= \sigma_{t+1} \frac{1}{1 - \alpha} \int_{F_Z^{-1}(\alpha)}^{\infty} \sqrt{a_1 y^2 + b} dF_Z(y),
\]
[B.3] in conjunction with
\[
\int_{F_Z^{-1}(\alpha)}^{\infty} \sqrt{a_1 y^2 + b} dF_Z(y) = \int_{\alpha}^{1} \sqrt{a_1 F_Z^{-1}(y)^2 + b} dy
\]
yields
\[
\text{AVaR}^\alpha_t(L_T) \geq \frac{1}{1 - \alpha} \int_{\alpha}^{1} F_Z^{-1}(u) du \left( \frac{1}{1 - \alpha} \int_{\alpha}^{1} \sqrt{a_1 F_Z^{-1}(y)^2 + b} dy \right)^{T-t-1} \sigma_{t+1} = \text{AVaR}^\alpha_t(L_T).
\]

\[\square\]

References

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. Mathematical Finance, 9(3):203–228, 1999.
A. A. Balkema and L. de Haan. Residual life time at great age. *Annals of Probability*, 2:792–804, 1974.

J. Bion-Nadal. Time consistent dynamic risk processes. *Stochastic Processes and Their Applications*, 119(2):633–654, 2009.

P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(01):137–162, 2011.

P. Cheridito and M. Stadje. Time-inconsistency of VaR and time-consistent alternatives. *Finance Research Letters*, 6(1):40–46, 2009.

K. Detlefsen and G. Scandolo. Conditional and dynamic convex risk measures. *Finance and Stochastics*, 9(4):539–561, 2005.

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. ISBN 3-540-60931-8.

H. Föllmer and T. Knispel. Entropic risk measures: Coherence vs. convexity, model ambiguity and robust large deviations. *Stochastics and Dynamics*, 11(02n03):333–351, 2011.

H. Föllmer and I. Penner. Convex risk measures and the dynamics of their penalty functions. *Statistics & Decisions*, 24(2006(1)):61–96, 2006.

H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter & Co., Berlin, extended edition, 2011. ISBN 978-3-11-021804-6.

C. Franq and J. Zakoian. *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley, Chichester, 2010.

A. McNeil and R. Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3):271–300, 2000.

A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005. ISBN 0-691-12255-5.

S. Peng. Nonlinear expectations, nonlinear evaluations and risk measures. In M. Frittelli and W. Runggaldier, editors, *Stochastic Methods in Finance*, pages 165–253. Springer, New York, 2004. Lecture Notes in Mathematics, vol. 1856.

I. Penner. *Dynamic Convex Risk Measures: Time Consistency, Prudence, and Sustainability*. PhD thesis, Humboldt-Universität zu Berlin, 2007.

J. Pickands. Statistical inference using extreme order statistics. *Annals of Statistics*, 3:119–131, 1975. ISSN 0090-5364.

S. Resnick. *Extreme Values, Regular Variation, and Point Processes*. Springer, New York, 1987.

F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and Their Applications*, 112(2):185–200, 2004.