LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Lehrstuhl für Statistik und ihre Anwendungen in Wirtschafts- und Sozialwissenschaften
Institut für Statistik
Prof. Dr. Kauermann

Master's Thesis

Non-parametric estimation of Range Value at Risk

Thomas Lux

Bearbeitungszeitraum:	05.10.2022 bis 03.03.2023
Betreuer:	Dennis Mao, M.Sc.
Verantw. Hochschullehrer:	Prof. Dr. Kauermann

Acknowledgments

I would like to appreciate the LaTeX-theses template by 'changkun' (https://github.com/changkun/lmu-thesis-latex).

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, 02.03.2023 . . . .

Contents

List of Figures

List of Tables

1 Introduction

Coming up with a reliable and universal way to quantify the risk associated with a financial position continues to challenge financial institutions, regulators and investors. The changes on the computation of minimum capital requirements for entities in the financial sector published in 1996 by the Basel Committee on Banking Supervision (BCBS 1996) propelled the adoption of 'Value at Risk' (VaR) as the 'standard' risk measure.
Today, VaR is popular in the financial industry and risk management research due to its conceptual simplicity (relative to other risk measures), applicability and efficient computation. Although popular, VaR has shortcomings, mainly (i) it lacks the sub-additive property, thus is not a coherent risk measure (as defined in Artzner et al. (1999)) and (ii) it is insensitive to potential losses exceeding the quantile of interest. Consequently, many alternative risk measures were introduced in order to overcome VaR's drawbacks. Acerbi & Tasche (2002) provide definitions of five risk measures that include losses beyond VaR:

- Conditional VaR (CVaR), as proposed in Rockafellar & Uryasev (2001)
- Expected Shortfall (ES), Acerbi & Tasche (2002)
- Tail Conditional Expectation (TCE), Artzner et al. (1999)
- Worst Conditional Expectation (WCE), Artzner et al. (1999)
- Spectral risk measures

Today, ES is probably the most discussed alternative, being both a coherent risk measure and sensitive to potential losses in the tail.
However, ES also has some drawbacks. ES is not elicitable, unlike VaR, meaning there is no consistent scoring function available (such as the squared loss, consistent for the mean). Though, ES is jointly elicitable with VaR, i.e. the tuple $(\mathrm{ES}, \mathrm{VaR})$ is elicitable (Section 2). Another favourable risk measure property is robustness against outliers and model misspecification, as data often deviates from underlying distribution assumptions. As coherent risk measures cannot be robust (Cont et al. 2010), ES is not robust, while VaR is (Section 2).
To obtain a robust risk measure that, unlike VaR, captures losses outside the specified quantile, Cont et al. (2010) introduced 'Range Value at Risk' (RVaR) as an alternative risk measure (Chapter 3). RVaR is not subadditive (coherent) and not directly elicitable, but like ES, jointly elicitable with VaR (Section 2). As RVaR is robust and includes VaR and ES as special cases, late research in financial mathematics, economics and statistics revolves around RVaR (e.g., Biswas & Sen (2023), Vasilev & Melnikov (2022), Fissler & Ziegel (2021), Bairakdar et al. (2020), Barendse (2020), Bernard et al. (2020), Embrechts et al. (2018), Li et al. (2018)).
Although recent research efforts around RVaR and its properties are considerable, few publications address the issue of estimating RVaR. Most recently, Biswas & Sen (2023) compares six non-parametric methods to estimate RVaR. A non-parametric approach is favourable as it (i) is model-free and thus, model robust; (ii) allows for a wide range of data dependency (often 'i.i.d.' assumption is not required) and lastly (iii) the tail portion of a loss distribution, where data is often sparse, is most relevant for practitioners and researchers, but complicates specification of a proper parametric loss model suited for the tail.

This theses aims to expand on the work of Biswas & Sen (2023), Brazauskas et al. (2008), Giannopoulos & Tunaru (2005), focusing on three non-parametric methods to estimate RVaR. We derive and compare the estimators' main theoretical properties as well as analyze their accuracy in an empirical and simulation study.

The theses is organized in six chapters. Chapter 1 motivates the research and scope of this paper. The 2nd Chapter reviews relevant literature and related publications. The 3rd Chapter outlines definitions and basic results on the risk measures in question. Chapter 4 defines three non-parametric estimation methods for RVaR and derives their main properties. Chapter 5 outlines the methodology and results of an empirical and simulation study, aiming to compare the estimators' behaviour and accuracy across different scenarios. The last Chapter (6) summarizes findings of this theses and suggests ideas for future work.

2 Literature review

This section reviews relevant literature on RVaR and its estimation, organized in reverse chronological order.

Biswas & Sen (2023)[1]

Objective: To describe six non-parametric estimators for RVaR and to analyse accuracy via simulation study and backtesting.
Method: Non-parametric estimators: Empirical estimator, Brazauskas et al.'s estimator, Kernel estimator, Tail-Trimmed estimator, Yamai and Yoshiba's estimator and Filtered Historical Method. ii) Simulation study: Monte-Carlo simulation ($30 \leq n \leq 1000$) to approximate mean squared error across five models: Generalized Pareto dist. ($\xi = 1/3$), Student's-t (df $= 4$), standard normal ($N(0,1)$) and two $\mathrm{ARMA}(1,1)$ with ($\phi = 0.95, \theta = -0.6, -0.9$).
Main conclusions: i) Adapts non-parametric estimators for ES to RVaR. ii) Although no estimator uniformly outperforms the other estimators in the simulation study, Brazauskas et al.'s estimator and the Filtered Historical Method perform best. iii) Simulation study and backtesting support recommendation of Filtered Historical Method to estimate RVaR except for heavy-tailed distribution (e.g., GPD).

Fissler & Ziegel (2021)

Objectives: i) To explore joint elicitability of RVaR and to characterise corresponding class of strictly consistent scoring functions. ii) To propose joint regression framework for RVaR and iii) to illustrate results via simulation study.
Method: i) Proof of joint elicitability of RVaR by characterising explicit class of scoring functions. ii) Comparison of predictive performance of estimators for triplet using Diebold-Mariano test ($n = 250, 10\,000$ repetitions).
Main conclusions: i) Proof that triplet of RVaR with two VaR components at different levels is elicitable. ii) Explicit representation of class of strictly consistent scoring functions for this triplet. iii) Proof that strictly consistent scoring function for triplet cannot be translation invariant or positively homogeneous. iv) Joint regression framework utilizing proposed scoring functions allows to substantially deviate from an i.i.d. assumption, while e.g. popular trimmed least squares method requires stationarity (a.o.).
Related publications: The notion of elicitability has been part of the debate about which risk measure is best, specifically within the discussion around backtesting (Ziegel 2014, Emmer et al. 2015, Davis 2016). While some argue that elicitability is not necessary for backtesting in practical applications (Acerbi & Szekely 2014, 2017, a.o.), Fissler & Ziegel (2016, 2021) demonstrate that elicitability is necessary for *comparative* backtesting. Classical backtesting aims to assess whether an estimation computed *ex ante* is sufficiently accurate w.r.t. observed *ex post* realizations. The goal is thus to assess the accuracy of the estimated value. In contrast, comparative backtesting uses the observed *ex post* sequence (e.g. sequence of portfolio returns) to provide a *ranking* of the different estimated values, hence 'comparative'.

[1] Initial results published April 2022 as conference paper in collection , extended results published February 2023 in 'Computation'.

Cont et al. (2010)

Objective: To introduce framework for studying the robustness of risk measurement procedures.
Method: i) Apply results from robust statistics to risk measures.
Main conclusions: i) Subadditive risk measure procedures cannot be robust. ii) Introduction of RVaR as a robust alternative to VaR and ES. iii) Robustness of risk measure procedure depends on both risk measure and estimation method.
Related publications: While VaR is not coherent (subadditive) in general, Embrechts et al. (2002) prove the subadditivity of VaR for elliptically distributed risks. Thus, under commonly used distribution models such as Normal, Student's t, Logistic or Laplace distribution, VaR is actually coherent. The findings of Cont et al. (2010) propelled the ongoing debate on which property is more favourable for financial risk measures, coherence or robustness. We point the reader to Cont, Deguest & Scandolo (2010), Kou, Peng & Heyde (2013) for arguments supporting robustness over coherence and Krätschmer, Schied & Zähle (2012, 2014, 2015) for an opposing view.

Brazauskas et al. (2008)

Objectives: i) To develop (non-) parametric estimator and corresponding confidence intervals for ES and ii) to compare their performance via simulation study.
Method: i) Non-parametric estimator based on empirical estimator and corresponding confidence intervals (point-wise and simultaneous). ii) Parametric estimator: Maximum likelihood and Delta-Method for Exponential, Pareto and Log-normal distributions and corresponding confidence intervals. iii) Simulation study via Monte-Carlo (n=10.000, 10 iterations) for three distributions (Exponential, Pareto and Log-normal).
Main conclusions: i) The proposed empirical estimator for ES is point-wise consistent, however not uniformly consistent over the interval $[0,1]$. ii) Derives closed form representation of point-wise and simultaneous confidence intervals for the empirical estimator and parametric estimators. iii) Coverage proportions of empirical confidence intervals are quite low in small samples but get reasonably close to intended 95% confidence level for $n \geq 250$. Although empirical estimator performs worse on average than parametric estimators, performance is similar across all distribution scenarios.

Giannopoulos & Tunaru (2005)

Objectives: i) To show that Filtered Historical Simulation (FHS) method can be used to estimate ES and ii) Resulting estimator is well-suited to forecast ES.
Method: i) Estimator assumes robust underlying volatility model (e.g. GARCH). ii) Empirical study: Daily NASDAQ100 returns for c. 5 years, fitting a modified GARCH$(1,1)$ model and applying FHS estimator.
Main conclusions: i) FHS method can be used to estimate ES and the resulting estimator is a coherent measure, even converging to a spectral risk measure. ii) Derives closed form asymptotic formula for standard errors for the ES estimators. iii) Empirical investigation confirms that ES estimates are more uncertain than VaR under FHS and that accuracy of estimator declines with increasing horizon.
Related publications: Dowd (2007) lists advantages of FHS applied to risk measure estimation: (i) merges upside of non-parametric historical simulation with advanced volatility

approach, accounting for market volatility conditions. (ii) fast computation, even for large portfolios. (iii) allows for e.g. VaR estimates exceeding the maximum observed historical loss. (iv) keeps correlation structure in return data without relying on variance-covariance matrices or conditional return distribution. (v) modifiable to account for autocorrelation and past cross-correlations in returns and to produce confidence interval estimates (for e.g. VaR) through combination with overall-survival or bootstrap approach. According to Zenti & Pallotta (2001), FHS is one of the best tools for calculating VaR.

3 Theoretical backdrop

The chapter is divided into two sections. The first section recollects main definitions and results on risk measures and defines VaR and ES. In the second, RVaR is formally introduced and its properties are compared to VaR and ES.

3.1 Risk measure basics

The following definitions and results are found in Yan (2018). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with real-valued random variables on Ω. We use $X \in \mathcal{G}$ to denote the *potential loss* (i.e. negative realizations of X represent a gain), F_X for its cumulative distribution function (CDF) and $\mathcal{G}$ for the set of potential losses to be considered. Further, let $\mathcal{D}$ denote the set of CDFs of these losses, e.g. $F_X \in \mathcal{D}$. We define risk measures and some of their properties closely following Yan (2018), Artzner et al. (1999).

Definition 3.1. A mapping $\rho : \mathcal{G} \mapsto \mathbb{R}$ is called *monetary risk measure* if it satisfies the following conditions for all $X, Y \in \mathcal{G}$:

- *Monotonicity*: If $X \leq Y$, then $\rho(X) \leq \rho(Y)$.
- *Translation invariance*: If $m \in \mathbb{R}$, then $\rho(X+m) = \rho(X) + m$.

Definition 3.2. A monetary risk measure is called *coherent*, if it satisfies monotonicity, translation invariance and for all $X, Y \in \mathcal{G}$:

- *Subadditivity*: It holds that $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
- *Positive homogeneity*: For all $\lambda \geq 0$, it holds that $\rho(\lambda X) = \lambda \cdot \rho(X)$.

We now turn to the statistical properties of risk measures. First, we introduce the notion of robustness for risk measures, which aims to determine how resilient the measure is towards changes in the underlying data. Intuitively, a robust risk measure should be unaffected by (small) errors in the model assumptions or outliers in the data. The concept of *qualitative robustness* proposed in Cont et al. (2010) naturally defines a risk measure as robust, if a small change in the distribution F results in a small change in the risk estimate:

Definition 3.3 (Def. 4, Cont et al. (2010))**.** A risk estimator ρ is *robust* at a loss distribution $F \in \mathcal{C}^2$ if for any $\varepsilon > 0$ there exists a $\delta > 0$ and $n_0 \geq 1$ such that for all $G \in \mathcal{C}^2$:

$$d(F,G) \leq \delta \quad \Longrightarrow \quad \forall n \geq n_0 : d(\mathcal{L}_n(\rho,\mathbb{P})), \mathcal{L}_n(\rho,\mathbb{Q})) \leq \varepsilon,$$

where d is the Lévy distance (Appendix A.1) and $\mathcal{L}_n(\rho, F)$ the distribution of risk measure ρ under CDF F.

Closely related is the *breakdown point*, quantifying the degree of robustness of a statistic. The breakdown point is “the smallest fraction of bad observations that may cause an estimator to take on arbitrarily large aberrant values”, thus the breakdown point of a non-robust statistic is 0, while the breakdown point is increasing with the degree of robustness

[2]The set of 'plausible' loss CDF's $\mathcal{C} \subseteq \mathcal{D}$ and assuming F is not an isolated point of $\mathcal{C}$ meaning there exists a loss distribution $G \in \mathcal{C}$ that is arbitrarily close to F.

of a statistic (Huber & Ronchetti 2009).
Second, we introduce the concepts of consistency and elicitability following Fissler & Ziegel (2016, 2021) before turning to VaR and ES.
Evaluating predictive performance is key when coming up with point forecasts ($\hat{R}_t$) for a statistical functional R (e.g. risk measure) of a quantity of interest y_t (e.g. portfolio returns). Predictive performance is generally measured in terms of the average realised score $1/n\sum_{t=1}^{n} S(\hat{R}_t, y_t)$ for a scoring or loss function S (the smaller the better). One aims to find a scoring function that favors the correct prediction (in the limit) or is *strictly consistent* (definition below). Prominent examples are (i) the squared loss $S(x,y) = (x-y)^2$ consistent for the mean or (ii) the absolute loss $S(x,y) = |x-y|$ consistent for the median. If there exists a consistent scoring function for the statistic (risk measure), it is called *elicitable*. The role of elicitability within financial risk management is debated (Section 2).

Definition 3.4 (Def. 2.1, Fissler & Ziegel (2016))**.** A *scoring function* is a $\mathcal{D}$-integrable function $S : \mathsf{A} \subseteq \mathbb{R}^k \times \Omega \mapsto \mathbb{R}$. We write $\bar{S}(x,F) := \int S(x,y)\,dF(y)$. It is said to be $\mathcal{D}$-*consistent* for a statistic $T : \mathcal{D} \mapsto \mathsf{A}$ if

$$\forall F \in \mathcal{D} \quad \forall x \in \mathsf{A}: \quad \bar{S}(T(F),F) \leq \bar{S}(x,F).$$

Furthermore, S is *strictly $\mathcal{D}$-consistent* for T if it is $\mathcal{D}$-consistent for T and if

$$\forall F \in \mathcal{D} \quad \forall x \in \mathsf{A}: \quad \bar{S}(T(F),F) = \bar{S}(x,F) \quad \Longrightarrow \quad x = T(F).$$

A functional $T : \mathcal{D} \mapsto \mathsf{A} \subseteq \mathbb{R}^k$ is called *k-elicitable*, if there exists a strictly $\mathcal{D}$-consistent scoring function for T (also referred to as *jointly* elicitable for $k \geq 2$).

We now formally introduce VaR and its properties.

Definition 3.5. For $p \in (0,1)$, the *Value at Risk (VaR) at level p* is defined as

$$\mathrm{VaR}_p(X) = \inf\{x \in \mathbb{R} : F_X(x) \geq p\}. \tag{3.1}$$

We omit the random variable in some situations and denote VaR_p for convenience. In this setting, VaR is the point such that $(1-p) \times 100\%$ of losses are on or below $\mathrm{VaR}_p(X)$ (Figure 3.1). We use the left-continuous version of VaR in this paper due to technical convenience. Moreover, we adopt the left-continuous inverse of F_X as *quantile function* (or F_X^{-1}), which therefore coincides with VaR.

As mentioned in the introduction, VaR is not a coherent risk measure as it lacks the subadditivity property. However, VaR satisfies monotonicity, translation invariance and positive homogeneity and is therefore a monetary risk measure (Appendix A.1). Furthermore, VaR is a robust monetary risk measure in the sense of Definition 3.3 with breakdown point $\min\{p, 1-p\}$ (Cont et al. 2010). At last, we introduce ES and highlight its properties.

Definition 3.6. For $p \in (0,1)$, the *Expected Shortfall (ES) at level p* is defined as

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_u(X)\,du. \tag{3.2}$$

This definition translates to the average loss in the worst $(1-p)\times 100\%$-cases. Again, we omit the random variable in most situations and denote ES_p. Another common representation of ES is

$$\mathrm{ES}_p(X) = \frac{\mathbb{E}\left[X; X \geq \mathrm{VaR}_p(X)\right]}{1-p} = \mathbb{E}\left[X | X \geq \mathrm{VaR}_p(X)\right], \tag{3.3}$$

which is equivalent to (3.2) for all $p \in [0,1]$ if F_X is continuous (Appendix A.1).

As discussed in the introduction, ES is not only a monetary risk measure, but also coherent (Appendix A.1). Because subadditive risk measures cannot be robust (Section 2), ES is not a robust risk measure (i.e. breakdown point of 0). Lastly, ES fails to be directly elicitable (Gneiting 2011), but is jointly elicitable, meaning the pair $(\mathrm{VaR}_p, \mathrm{ES}_p) : \mathcal{F} \mapsto \mathbb{R}^2$ is elicitable for any $p \in (0,1)$ subject to mild conditions on $\mathcal{F}$ (Fissler & Ziegel 2016).

3.2 Range Value at Risk

In this section, we formally introduce RVaR and highlight its properties. While discussing statistical robustness of monetary risk measures, Cont et al. (2010) introduce RVaR as a compromise between VaR and ES (Section 2).

Definition 3.7. The *Range Value at Risk (RVaR) at levels* $0 < p < q < 1$ of a loss X is defined as

$$\mathrm{RVaR}_{p,q}(X) = \frac{1}{q-p}\int_p^q \mathrm{VaR}_u(X)\, du. \tag{3.4}$$

That is, $\mathrm{RVaR}_{p,q}$ is the average of all VaR_u with u between p and q (Figure 3.1). We omit the random variable when appropriate, abbreviating with $\mathrm{RVaR}_{p,q}$. ES represents a special case of RVaR, as

$$\mathrm{RVaR}_{p,1}(X) := \lim_{q\uparrow 1} \mathrm{RVaR}_{p,q}(X) = \mathrm{ES}_p(X).$$

Further, we set RVaR as VaR if p and q coincide, i.e.

$$\forall 0 < p = q < 1 : \quad \mathrm{RVaR}_{p,q}(X) := \mathrm{VaR}_p(X).$$

The definition yields a useful inequality:

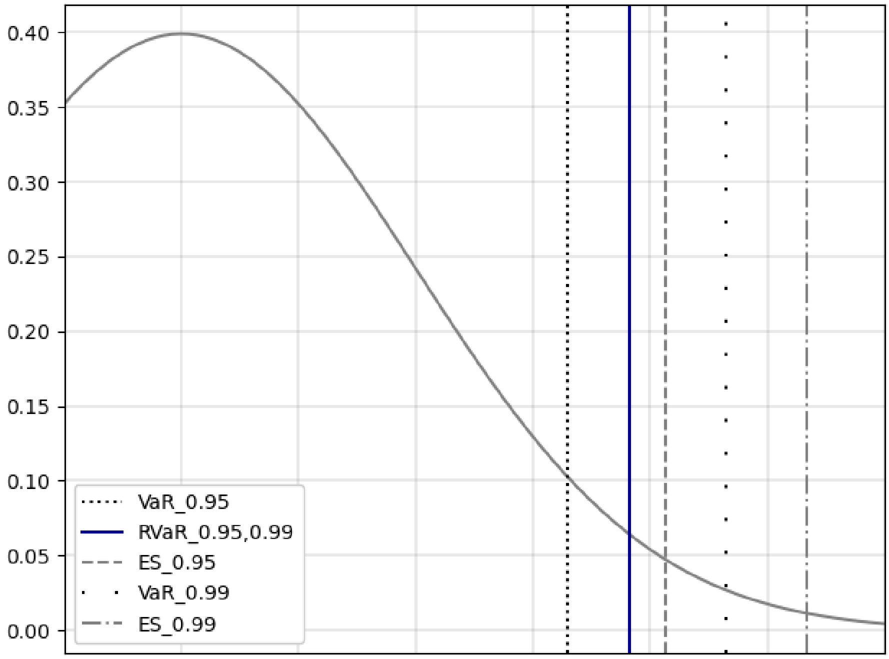
$$\mathrm{VaR}_p(X) \leq \mathrm{RVaR}_{p,q}(X) \leq \mathrm{VaR}_q(X), \tag{3.5}$$

which holds for any loss X and $0 < p \leq q \leq 1$, illustrated in Figure 3.1.

We can also express RVaR in terms of ES, which is particularly convenient when deriving RVaR estimators in Chapter 4. We have for any $0 < p < q < 1$,

$$\begin{aligned}
\mathrm{RVaR}_{p,q} &= \frac{1}{q-p}\left(\int_p^1 \mathrm{VaR}_u\, du + \int_1^q \mathrm{VaR}_m\, dm\right) \\
&= \frac{1}{q-p}\left(\frac{1-p}{1-p}\int_p^1 \mathrm{VaR}_u\, du - \frac{1-q}{1-q}\int_q^1 \mathrm{VaR}_m\, dm\right) \\
&= \frac{(1-p)\mathrm{ES}_p \; - \; (1-q)\mathrm{ES}_q}{q-p}.
\end{aligned} \tag{3.6}$$

Figure 3.1: PDF of $\mathcal{N}(0,1)$ incl. VaR, ES and $\mathrm{RVaR}_{p,q}$ $(p = 0.95,\ q = 0.99)$.



Like VaR, RVaR is a monetary risk measure but not coherent, excluding the special case $\mathrm{RVaR}_{p,1} = \mathrm{ES}_p$ here and in the following (see Appendix A.1).
RVaR is, unlike ES, a robust risk measure (Definition 3.3). Analyzing $\mathrm{RVaR}_{p,q}$'s robustness, Fissler & Ziegel (2021) conclude a breakdown point of $\min\{p, 1-q\}$. Thus, $\mathrm{RVaR}_{p,q}$ is less robust than VaR_p (breakdown point of $\min\{p, 1-p\}$) and more robust than ES_p with breakdown point of 0.
Like ES, RVaR fails to be directly elicitable (Wang & Wei 2020), but is jointly elicitable as Fissler & Ziegel (2021) show that the triplet $(\mathrm{VaR}_p, \mathrm{VaR}_q, \mathrm{RVaR}_{p,q})$ is elicitable under weak regularity conditions (Section 2).

4 Estimators for Range Value at Risk

Biswas & Sen (2023) evaluate six non-parametric methods for estimating RVaR, namely the Empirical estimator (van der Vaart 2012), Brazauskas et al.'s estimator (Brazauskas et al. 2008), Filtered Historical Method (initially applied to ES by Giannopoulos & Tunaru (2005)), Kernel Estimator (Biswas & Sen (2019), leveraging Altman & Léger (1995)), Tail-Trimmed estimator (Hill 2013) and Yamai and Yoshiba's estimator (Yamai & Yoshiba 2002).

Although no estimator uniformly outperforms the other estimators in Biswas & Sen (2023) simulation study, Brazauskas et al.'s Estimator and the Filtered Historical Method perform best. Therefore we focus on these estimators along with the 'natural' Empirical estimator in the subsequent sections and our study.

4.1 Empirical estimator

Let $X_1,\dots,X_n$ be a random sample of observed losses from a CDF $F_X \in \mathcal{D}$ and $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ be the order statistics of the random sample. We denote the empirical distribution function by $\hat{F}(x)$ (Section A.2.1).

Following (van der Vaart 2012, Ch. 21.2), the *empirical quantile function* is defined as

$$\hat{F}^{-1}(p) = X_{(i)}, \quad \text{for } p \in \left(\frac{i-1}{n}, \frac{i}{n}\right]. \tag{4.1}$$

This estimator is reasonable for F^{-1} (or VaR in our case), especially if the underlying distribution is unknown. The following theorem ensures weak convergence (also known as convergence in distribution) of the quantile estimator towards the true quantile.

Theorem 4.1 (Lemma 21.2, van der Vaart (2012))**.** *For any sequence of cumulative distribution functions, $F_n^{-1} \rightsquigarrow F^{-1}$ if and only if $F_n \rightsquigarrow F$, where '$\rightsquigarrow$' refers to weak convergence, i.e. $F_n^{-1}(x) \xrightarrow{d} F^{-1}(x)$ at every x where F^{-1} is continuous.*

Proof (van der Vaart 2012). Let $F_n^{-1} \rightsquigarrow F^{-1}$ and let $U \sim \mathrm{U}[0,1]$. Because F_n^{-1} has at most countable many discontinuity points, we have $F_n^{-1}(U) \xrightarrow{a.s.} F^{-1}(U)$. Consequently, $F_n^{-1}(U)$ converges weakly towards $F^{-1}(U)$. This implies $F_n \rightsquigarrow F$ by the quantile transformation property (Lemma A.1).

Now let $F_n \rightsquigarrow F$ and let $V \sim \mathcal{N}(0,1)$ with CDF Φ. Hence $F_n(V) \xrightarrow{\text{a.s.}} F(V)$ because convergence can only fail at discontinuity points of F and V is continuous. Further, $\Phi\left(F_n^{-1}(x)\right) = \mathbb{P}(F_n(V) < x)$ (Lemma A.2, (i)) converges to $\mathbb{P}(F(V) < x) = \Phi\left(F^{-1}(x)\right)$ at every x at which the limit function is continuous, especially at every x at which F^{-1} is continuous. By the continuity of Φ^{-1}, $F_n^{-1}(x) \xrightarrow{a.s.} F^{-1}(x)$ for every such x. □

As the empirical estimator $\hat{F}(x)$ satisfies weak consistency for F (Equation (A.4)), the sequence of estimators $\hat{F}_n^{-1}$ is therefore weakly consistent for the quantile function F^{-1}. With weak consistency established, we explore the quantile estimator's asymptotic properties.

Theorem 4.2 (Quantile estimator is asymptotically normal; Cor. 21.5, (van der Vaart 2012))**.** *Fix $0 < p < 1$. If F is differentiable at the p-th quantile $F^{-1}(p)$ with positive* *derivative* $f\left(F^{-1}(p)\right)$*, then the quantile estimator* $\hat{F}^{-1}(p)$ *is asymptotically normal:*

$$\sqrt{n}\left(\hat{F}^{-1}(p)-F^{-1}(p)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{p(1-p)}{f^2(F^{-1}(p))}\right). \tag{4.2}$$

Proof (Stephens 2006). We establish in the Appendix (A.3) that $\mathbb{E}\left[\hat{F}(x)\right]=F(x)$. Also, one can easily verify that $\mathbb{V}\text{ar}\left[\hat{F}(x)\right]=F(x)(1-F(x))$. By the central limit theorem, we have

$$\sqrt{n}\left(\hat{F}(x)-F(x)\right) \xrightarrow{d} \mathcal{N}\left(0, F(x)[1-F(x)]\right).$$

Now consider the transformation through $g(t)$ for $0<t<1$ by $g(t):=F^{-1}(t)$. The first derivative of g is

$$g'(t)=\frac{d}{dt}\left[F^{-1}(t)\right]=\frac{1}{f(F^{-1}(t))},$$

as

$$\begin{aligned} y=F^{-1}(t) \quad &\iff \quad F(y)=t \quad \implies \quad f(y)dy=dt \\ &\implies \quad \frac{dy}{dt}=\frac{1}{f(y)}=\frac{1}{f(F^{-1}(t))}. \end{aligned}$$

Then, using the Delta-Method (Lemma A.3), we get

$$\sqrt{n}\left(F^{-1}(\hat{F}(x))-F^{-1}(F(x))\right) \xrightarrow{d} \mathcal{N}\left(0, F(x)[1-F(x)]\frac{1}{\left(f[F^{-1}(F(x))]\right)^2}\right).$$

Now let F be differentiable at the p-th quantile $F^{-1}(p)$ with positive derivative $f\left(F^{-1}(p)\right)$ and set $p=F(x)$. This yields

$$\sqrt{n}\left(F^{-1}(\hat{F}(x))-F^{-1}(p)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{p(1-p)}{f^2(F^{-1}(p))}\right).$$

Lastly, $F^{-1}(\hat{F}(x))$ is a random variable that lies between the $(p-1)$st and the pth sample quantile, thus converges almost surely to $\hat{F}^{-1}(p)=X_{(i)}$ with $p\in((i-1)/n, i/n]$ and therefore in distribution. This concludes (4.2) and the proof. □

The asymptotic variance of the empirical quantile behaves as expected, as the numerator gets smaller for more extreme quantiles, i.e. more information/data is needed to predict the tail quantiles in comparison to the central quantiles. The denominator on the other hand is largest when the CDF F is alternating most around the pth quantile, i.e. there is a higher probability for samples around $F^{-1}(p)$, thus estimation is more accurate.

Knowing that the empirical quantile is not only a reasonable, but also a consistent and asymptotically normal estimator for VaR, we can confidently utilize it to estimate ES. Recall from definition (3.2) that ES is the average loss in the worst $(1-p)\times 100\%$ cases, thus the empirical estimator of ES is naturally defined as the weighted average of excessive losses larger than $\hat{F}^{-1}(p)=X_{([np]+1)}$, i.e.

$$\text{EMP}_p(X)=\frac{1}{\lfloor n(1-p)\rfloor+1}\sum_{i=1}^{n} X_i \mathbf{I}_{\{X_i\geq \hat{F}^{-1}(p)\}}, \tag{4.3}$$

where $\lfloor\cdot\rfloor$ is the integer floor function.
Given the asymptotic and consistence properties of the empirical quantile function, it is not surprising that EMP is not only consistent for ES, but also asymptotically normal, ensured by the following theorem.

Theorem 4.3 (Theorem 1 in Chen (2007)). *Let $l \in (0,1)$ such that $\sup_{A\in\mathcal{F}_1^i, B\in\mathcal{F}_{i+k}^\infty} |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)| \leq Cl^k$ for all $k \geq 1$ and a positive constant C. Let the stationary CDF F of the stationary process $\{X_t\}$ be absolutely continuous and its PDF f have continuous second derivatives in a neighbourhood of VaR_p. Also, let F_k be the joint CDF's of (X_1, X_{k+1}) for $k \geq 1$ with bounded second partial derivatives in a neighbourhood of VaR_p. Lastly, let $\mathbb{E}\left[|X_t|^{2+\delta}\right] < \infty$ for some $\delta > 0$. Then,*

$$(EMP_p(X) - ES_p(X)) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \frac{\sigma_0^2(p;n)}{n(1-p)^2}\right), \tag{4.4}$$

where

$$\sigma_0^2 := \mathbb{V}\mathrm{ar}\left[(X_1 - VaR_p)_+ \; + 2\sum_{k=1}^{n-1} \mathbb{C}\mathrm{ov}\left[(X_1 - VaR_p)_+; (X_{k+1} - VaR_p)_+\right]\right], \tag{4.5}$$

and $(a-b)_+ := (a-b)\boldsymbol{I}_{\{a\geq b\}}$.

Proof. We refer the reader to Chen (2007) for a comprehensive proof, but outline the main arguments here.
First, the author expresses the empirical estimator as a linear combination of VaR, ES, empirical distribution functions and a landau term. Next, he employs the blocking technique, re-writing the estimator as a sum of three blocks. Two of three blocks converge to zero, which is shown directly. To establish the asymptotic behaviour of the remaining block, the author employs Bradley's lemma (Bosq 1998), expressing the block as a stationary α-mixing sequence[3](Rosenblatt 1956, a.o.). This allows for applying moment bounds given in Yokoyama (1980), which ensures the Liapounov condition for the central limit theorem of triangular arrays and together, the convergence of the last block. Applying Slutsky's theorem (van der Vaart 2012, Lemma 2.8) permits combining the three blocks in limit and completes the proof. □

Note the assumption of an absolutely continuous CDF F_X, which is not needed to establish asymptotic normality of the Brazauskas et al.'s estimator. More on this in the subsequent section 4.2.
Furthermore, as $(1-p)$ is typically small (usually $1-5\%$ in financial risk management), EMP's asymptotic variance can be large, especially for smaller n's. This poses a challenge for statistical inference utilizing the EMP estimator. Nevertheless, the empirical estimator is commonly used in practice for ES forecasting (e.g. Wu et al. (2021), Mehlitz & Auer (2020)).
At last, we can conclude an asymptotically normal, hence consistent estimator for $\mathrm{RVaR}_{p,q}$ leveraging the identity in (3.6):

$$\mathrm{EMP}_{(p,q)}(X) = \frac{(1-p)\mathrm{EMP}_p(X) - (1-q)\mathrm{EMP}_q(X)}{q-p}. \tag{4.6}$$

Implementation

Implementation of Empirical estimator for Chapter 5 is straightforward, utilizing `pandas series' sort_values` and `sum` methods to derive EMP_p and EMP_q as in Equation (4.3). Then, Identity (4.6) yields the RVaR estimate.

[3]Per Chen (2007), the dependence described by α-mixing is the weakest as it is implied by other types of mixing. Refer to Doukhan (2012) for a comprehensive review of different stochastic mixing notions.

4.2 Brazauskas et al.'s estimator

Brazauskas et al. (2008) propose an estimator for ES (or conditional tail expectation in their paper) related to the empirical estimator. Utilizing the definition of ES in (3.2) and replacing 'true' VaR with the empirical quantile estimator we get Brazauskas et al.'s estimator, i.e.

$$\mathrm{BRZ}_p(X) = \frac{1}{1-p}\int_p^1 \hat{F}^{-1}(u)\,du. \tag{4.7}$$

The remainder of the subsection explores BRZ's asymptotic properties and establishes it's consistency, following the arguments in Brazauskas et al. (2008).
We begin by showing that BRZ_p converges to ES_p for every fixed $p \in [0,1]$ almost surely.

Theorem 4.4 (Theorem 2.1 in Brazauskas et al. (2008))**.** *Assuming that the first moment $\mathbb{E}[X]$ is finite, we have for every $p \in [0,1]$ that $BRZ_p(X)$ converges to $ES_p(X)$ almost surely and thus in probability. In other words, $BRZ_p(X)$ is a strongly (and thus weakly) consistent estimator of $ES_p(X)$ for every fixed $p \in [0,1]$.*

Proof (p. 3601, Brazauskas et al. (2008)). Assume $\mathbb{E}[X] < \infty$. Theorem 4.4 can be simplified as

$$\begin{aligned}
\forall p \in [0,1]: \quad & \frac{1}{1-p}\int_p^1 \hat{F}^{-1}(u)\,du \quad \xrightarrow{a.s} \quad \frac{1}{1-p}\int_p^1 \mathrm{VaR}_u(X)\,du \quad \Longleftrightarrow \\
\forall p \in [0,1]: \quad & \int_p^1 \hat{F}^{-1}(u)\,du \quad \xrightarrow{a.s.} \quad \int_p^1 \mathrm{VaR}_u(X).
\end{aligned}$$

The preceding convergence follows if the L_1-distance of $\hat{F}^{-1}(u)$ and VaR_u converges almost surely to 0, i.e. if

$$\int_0^1 |\hat{F}^{-1}(u) - \mathrm{VaR}_u|\,du \xrightarrow{a.s.} 0, \tag{4.8}$$

holds. The two following conditions are equivalent to (4.8) (Lemma A.4):

(i) $F_n \xrightarrow[n\to\infty]{d} F$,

(ii) $\int |x|\,dF_n \xrightarrow[n\to\infty]{a.s.} \int |x|\,dF$.

The first statement follows directly from the Glivenko-Cantelli theorem (Theorem A.1), while the second statement is equal to

$$\frac{1}{n}\sum_{i=1}^{n} |X_i| \xrightarrow[n\to\infty]{a.s.} \mathbb{E}[|X|].$$

This of course holds true by the strong law of large numbers for $|X|$ as $\mathbb{E}[|X|] < \infty$ by assumption, which concludes the proof . □

With point-wise consistency established, we briefly illustrate why uniform consistency across the interval $[0,1]$ does not hold, i.e.

$$\sup_{0<p<1} |\mathrm{BRZ}_p(X) - \mathrm{ES}_p(X)| > 0. \tag{4.9}$$

As $\mathrm{BRZ}_p(X) \leq X_{(n)}$ by definition, the following holds:

$$\begin{aligned}\sup_{0<p<1} |\mathrm{BRZ}_p(X) - \mathrm{ES}_p(x)| &\geq \sup_{1-1/n<p<1} |\mathrm{BRZ}_p(X) - \mathrm{ES}_p(X)| \\ &= \sup_{1-1/n<p<1} |X_{(n)} - \mathrm{ES}_p(x)| \\ &\geq - \sup_{1-1/n<p<1} X_{(n)} + \sup_{1-1/n<p<1} \mathrm{ES}_p(X), \end{aligned} \tag{4.10}$$

where the last inequality holds by the reverse triangle inequality. The right-hand side of (4.10) is infinite: While the order statistic is finite (almost surely), the supremum is infinite for distributions with infinite right-end support (e.g. normal or Student's-t distribution). Therefore, the right-hand side of (4.10) does not converge (is greater than zero), and thus, in turn, the left-hand side of (4.10) does not converge either. This concludes that BRZ_p is not a uniformly consistent estimator of ES_p across the interval $[0,1]$.
With BRZ's consistency properties established, we explore its asymptotic behavior. In fact, BRZ is an asymptotically normal estimator for ES by the following theorem.

Theorem 4.5 (Theorem 3.1 in Brazauskas et al.)**.** *Assume that the second moment $\mathbb{E}\left[X^2\right]$ is finite. Let $p \in [0,1]$ be fixed, and let the CDF F_X be continuous at the point $F_X^{-1}(p)$. Then*

$$\sqrt{n}\left(BRZ_p(X) - ES_p(X)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_X^2(p)\right), \tag{4.11}$$

with the variance

$$\sigma_X^2(p) = \frac{1}{(1-p)^2} \int_{F_X^{-1}(p)}^{\infty} \int_{F_X^{-1}(p)}^{\infty} \left(F_X(x \wedge y) - F_X(x) F_X(y)\right) dx\, dy. \tag{4.12}$$

In particular, statement (4.11) holds for any fixed $p \in [0,1]$ if the CDF F_X is continuous across $\mathbb{R}$.

Proof (p. 3602, (Brazauskas et al. 2008)). By definition, we have

$$\mathrm{BRZ}_p(X) - \mathrm{ES}_p(X) = \frac{1}{1-p} \int_p^1 \left(\hat{F}^{-1}(u) - F^{-1}(u)\right) du. \tag{4.13}$$

In the following, we extract a sum of random variables from the integral in (4.13). To do this, we focus on the integral below for now, omitting the factor:

$$\int_p^1 \left(\hat{F}^{-1}(u) - F^{-1}(u)\right) du. \tag{4.14}$$

Note that (4.14) is the difference between the empirical mean $\overline{X}$ and the expected value $\mathbb{E}[X]$ iff $p = 0$, as

$$\int_0^1 F^{-1}(u)\, du \overset{u=F(x)}{=} \int_{-\infty}^{\infty} F^{-1}(F(x)) \frac{f(x)dx}{dx}\, dx = \int_{-\infty}^{\infty} x f(x)\, dx = \mathbb{E}[X].$$

Moreover, we can express the area between $\hat{F}^{-1}$ and F^{-1} in terms of the CDF and empirical CDF for $p = 0$,

$$\int_0^1 \left(\hat{F}^{-1}(u) - F^{-1}(u)\right) du = - \int_{-\infty}^{\infty} \left(\hat{F}(x) - F(x)\right) dx. \tag{4.15}$$

We utilize this identity to extract a sum of i.i.d. random variables from the integral. We begin with the following statement:

$$\int_p^1 \left(\hat{F}^{-1}(u) - F^{-1}(u)\right) du = -\int_{F^{-1}(p)}^{\infty} \left(\hat{F}(x) - F(x)\right) dx \; + \; \hat{R}(p), \tag{4.16}$$

where $\hat{R}(p)$ is defined by Equation (4.16) itself. Iff $p = 0$, $\hat{R}(p)$ equals 0 as per Equation (4.15). Next, we define

$$-\hat{V}(p) \; := \; \hat{R}(p), \tag{4.17}$$

where

$$\begin{aligned}\hat{V}(p) = -\hat{R}(p) &= -\int_p^1 \left(\hat{F}^{-1}(u) - F^{-1}(u)\right) du - \int_{F^{-1}(p)}^{\infty} \left(\hat{F}(x) - F(x)\right) dx \\ &\overset{(4.15)}{=} \int_0^p \left(\hat{F}^{-1}(u) - F^{-1}(u)\right) du + \int_0^{F^{-1}(p)} \left(\hat{F}(x) - F(x)\right) dx.\end{aligned}$$

The reasoning behind these steps is that the process $\hat{V}(p)$ is well-researched, known as the (general) Vervaat process. In fact, $\hat{V}(p)$ is non-negative for all $p \in [0,1]$ (Lemma A.5) and satisfies the following upper-bound:

$$\hat{V}(p) \leq -\left(\hat{F}(F^{-1}(p)) - p\right)\left(\hat{F}^{-1}(p) - F^{-1}(p)\right), \tag{4.18}$$

for any CDF F (Lemma A.5). As we assume continuity of F at the point $F^{-1}(p)$, we of course have $p = F(F^{-1}(p))$ and the even stronger upper-bound:

$$\begin{aligned}|\hat{V}(p)| &\leq |\hat{F}(F^{-1}(p)) - F(F^{-1}(p))| \cdot |\hat{F}^{-1}(p) - F^{-1}(p)| \\ &\leq \sup_{x \in \mathbb{R}} |\hat{F}(x) - F(x)| \cdot |\hat{F}^{-1}(p) - F^{-1}(p)|.\end{aligned} \tag{4.19}$$

The supremum term in (4.19) is of the order $\mathcal{O}(\sqrt{n})$ by the Kolmogorov-Smirnov theorem (e.g. Shorack & Wellner (2009, p. 142)). Thus, if $\hat{F}^{-1}(p) \xrightarrow{P} F^{-1}(p)$, we have $\sqrt{n}|\hat{V}(p)| = o(1)$. As we assume continuity of F at $F^{-1}(p)$, we have $\hat{F}^{-1}(p) \xrightarrow{P} F^{-1}(p)$ by Theorem 4.1. Applying these steps to (4.13) yields for every fixed $p \in [0,1]$,

$$\begin{aligned}\sqrt{n}(\mathrm{BRZ}(p) - \mathrm{ES}(p)) &= -\frac{\sqrt{n}}{1-p}\int_{F^{-1}(p)}^{\infty} \left(\hat{F}(x) - F(x)\right) dx - \frac{\sqrt{n}}{1-p}\hat{V}(p) \quad \text{by (4.16), (4.17)} \\ &= -\frac{\sqrt{n}}{1-p}\int_{F^{-1}(p)}^{\infty} \left(\hat{F}(x) - F(x)\right) dx + o(1) \\ &= -\frac{1}{1-p}\frac{\sqrt{n}}{n}\int_{F^{-1}(p)}^{\infty} \sum_{i=1}^{n} \mathbf{I}_{\{X_i \leq x\}} - F(x)\, dx + o(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} H_p(X_i) + o(1),\end{aligned}$$

where

$$H_p(X_i) := -\frac{1}{1-p}\int_{F^{-1}(p)}^{\infty} \left(\mathbf{I}_{\{X_i \leq x\}} - F(x)\right) dx. \tag{4.20}$$

For every fixed $p \in [0,1]$, the random variables $H_p(X_i)$ are centered and i.i.d. (for all $1 \leq i \leq n$), as

$$\mathbb{E}[H_p(X_i)] = -\frac{1}{1-p}\int_{F^{-1}(p)}^{\infty} \left(\mathbb{E}\left[\mathbf{I}_{\{X_i \leq x\}}\right] - F(x)\right) dx = 0.$$

Moreover, $H_p(X_i)$ has variance $\sigma_X^2(p)$ defined in (4.12), as

$$\begin{aligned}
\mathbb{V}\text{ar}[H_p(X_i)] &= \mathbb{E}\left[H_p(X_i)^2\right] - 0 = \frac{1}{(1-p)^2}\mathbb{E}\left[\left(\int_{F^{-1}(p)}^{\infty}\left(\mathbf{I}_{\{X_i \leq x\}} - F(x)\right) dx\right)^2\right] \\
&= \frac{1}{(1-p)^2}\mathbb{E}\left[\int_{F^{-1}(p)}^{\infty}\int_{F^{-1}(p)}^{\infty} \mathbf{I}_{\{X_i \leq x \wedge y\}} - \mathbf{I}_{\{X_i \leq x\}}F(y) - \mathbf{I}_{\{X_i \leq y\}}F(x) + F(x)F(y)\,dx\,dy\right] \\
&= \frac{1}{(1-p)^2}\int_{F^{-1}(p)}^{\infty}\int_{F^{-1}(p)}^{\infty} F(x \wedge y) - 2F(x)F(y) + F(x)F(y)\,dx\,dy = \sigma_X^2(p).
\end{aligned}$$

The variance $\sigma_X^2(p)$ is finite for every $p \in [0,1]$, iff $\mathbb{E}\left[X^2\right] < \infty$ which holds true by assumption. The central limit theorem concludes (4.11) and the proof. □

As with the empirical estimator, we can formulate a point-wise consistent and asymptotically normal non-parametric estimator for $\text{RVaR}_{p,q}$:

$$\text{BRZ}_{(p,q)}(X) = \frac{(1-p)\text{BRZ}_p(X) - (1-q)\text{BRZ}_q(X)}{q-p}. \tag{4.21}$$

Unlike the empirical estimator, we do not assume continuity of F for the consistency argument, and merely continuity in $F_X^{-1}(p)$ for the asymptotic argument (vs. uniform continuity assumption for asymptotic normality of EMP). This makes the BRZ estimator potentially more robust when dealing with (partially) discontinuous distributions.

Implementation

As the integral in Definition (4.7) reduces to a finite sum in our empirical and numerical study setting, implementation of Brazauskas et al.'s estimator is analogous to Empirical estimator implementation (Section 4.1).

4.3 Filtered Historical Method

The Filtered Historical Method (FHS) was introduced by Barone-Adesi et al. (1998, 1999) and initially applied to ES estimation by Giannopoulos & Tunaru (2005) (Section 2). FHS re-samples historical returns within a conditional volatility framework (e.g. GARCH). The main difference to other non-parametric approaches is, that returns are adjusted by the corresponding conditional volatility first. Then, the standardised returns are scaled by the volatility reflecting current or future conditions (expectations).

FHS method can be implemented in four steps (Giannopoulos & Tunaru 2005):

(i) Fit conditional volatility model to historical data (e.g. Biswas & Sen (2023) use a GARCH$(1,1)$).

(ii) Generate daily volatilities for each day (or any other time period) of sample period and standardise realised returns (dividing by corresponding volatility σ):

$$\widetilde{X}_i = \frac{X_i}{\sigma^X_{t-i}}, \quad \text{for } i = 1, \ldots, n \tag{4.22}$$

Standardised values should be i.i.d.

(iii) Bootstrap with replacement from set of standardised returns $(\widetilde{X}_1, \ldots, \widetilde{X}_n)$ a large number L of drawings $(\check{X}_1, \ldots, \check{X}_L)$. Then, each drawing is multiplied by the volatility forecast at the horizon σ_X to obtain a sample of i.i.d. values, as large as needed, from the returns that will occur on that day yielding:

$$\left((\sigma_X \check{X}_1), \ldots, (\sigma_X \check{X}_L)\right) \tag{4.23}$$

Note that due to the bootstrapping procedure drawing with replacement, $\check{X}_j$ can take any of the values $(\widetilde{X}_1, \ldots, \widetilde{X}_n)$ for $j = 1, \ldots, L$.

(iv) The estimator for ES is then:

$$\text{FHS}_p(X) = \frac{\sigma_X}{\theta} \sum_{i=1}^{\theta} \check{X}_{(i)}, \tag{4.24}$$

where $\theta = \lfloor pL \rfloor$ and $\lfloor \cdot \rfloor$ the integer floor function.

Consequently, the estimator of $\text{RVaR}_{p,q}$ is:

$$\text{FHS}_{(p,q)}(X) = \frac{(1-p)\text{FHS}_p(X) - (1-q)\text{FHS}_q(X)}{q-p} \tag{4.25}$$

Giannopoulos & Tunaru (2005) not only show that the FHS method can be used to estimate ES, but also that the resulting estimator maintains the coherence property of ES. Further, a closed form asymptotic formula for the variance of this estimator is derived.

Implementation

We implement FHS estimator for Chapter 5 closely following steps (i)-(iv) above (step (iii) is omitted for Monte-Carlo simulations).
For (i), we use a GARCH$(1,1)$ model with normally distributed errors via the `arch` package. In step (ii), we standardize returns using its `conditional_volatility` method. Step (iii) generates one-step ahead forecasts with `arch's fit` method. Note that the volatility forecast method has significant impact on RVaR estimates with one-step ahead forecasts resulting in fairly volatile rolling RVaR estimates (Section 5.2.1). Bootstrapping is achieved via `pandas series' sample` method, using $L = 1\,000$ draws. After testing various values for L, we consider $L = 1\,000$ a reasonable trade-off between computational load and robustness of estimates. Step (iv) is implemented similar to the Empirical estimator (Section 4.1).

5 Empirical & Simulation Study

5.1 Methodology

The subsequent sections explain the methods used in the empirical and simulation study.

5.1.1 Empirical study

We conduct all computational implementations using the `Python` programming language, making use of the following packages: `numpy`, `pandas`, `datetime`, `yfinance`, `arch`, `matplotlip`. We use daily price data of the S&P 500 and DAX stock indices from Yahoo!finance from January 1985 and January 1988 to December 2022, respectively. We consider the daily log-loss, calculated as $X_t = \log(P_{t-1}/P_t)$ where P_t is the closing price on day t. We use significance levels $\alpha, \beta \in \{1\%, 2.5\%, 5\%\}$, implying the following pairs for p and q: $\{(0.95, 0.975), (0.95, 0.99), (0.975, 0.99)\}$ as $p < q$ for RVaR. Significance levels $\alpha, \beta \in \{1\%, 5\%\}$ are common in risk management research, while BCBS recommends significance level of 2.5% for ES (BCBS 2013). We employ a rolling estimation window of 250, 500 and 1 000 daily observations, same as the sample size of the MC simulation in 5.1.2. The study focuses on the RVaR estimators described in Chapter 4, i.e. the Empirical estimator, Brazauskas et al.'s estimator and Filtered Historical Simulation (FHS) estimator. In addition to the estimates, we compute 99% confidence bands using the estimators' observed standard deviation as there are no closed form variance formulas available for the three RVaR estimators in question (Appendix A.6).

5.1.2 Numerical procedures

As in the empirical study, implementations use `Python`, utilizing `scipy:stats` in addition to the packages mentioned in Section 5.1.1. We use Monte Carlo (MC) simulation to evaluate accuracy of the three RVaR estimators introduced in Chapter 4, employing Normal and Student's-t distribution model as well as a GARCH$(1,1)$ model with constant mean and Gaussian residuals. Model parameters are estimated via Maximum-Likelihood estimation utilizing the S&P 500 data described in Section 5.1.1. The significance levels used are the same as in the empirical study. We employ 1 000 MC draws, which provides sufficient results (e.g. Yi et al. (2014)) and consider 250, 500 and 1 000 samples [4], equal to the rolling estimation window size in Section 5.1.1. To evaluate accuracy, we use Bias, Relative Bias (R. Bias), Root-Mean-Squared-Error (RMSE) and the variance of the estimate (σ^2), see Appendix A.3 for definitions.

5.2 Results

The empirical and numerical simulation study confirm that all three estimators examined in Chapter 4 are suitable for RVaR estimation. Empirical and Brazauskas et al.'s estimator behave similarly across both studies as the two estimators are based on empirical quantiles alike. FHS estimates in the empirical study are more volatile compared to Empirical and Brazauskas et al.'s estimates, due to the one-step ahead volatility forecasting (Section 4.3).

[4] While 250 is the minimum sample size recommended by the BCBS, sample size of 500 and 1 000 is common in risk management literature.

At the same time, FHS estimator accuracy in the simulation study is competitive in most scenarios.

5.2.1 Empirical results

This subsection presents results from applying the different estimation methods outlined in Chapter 4 to S&P 500 and DAX log-loss data. Table 5.1 outlines descriptive statistics of S&P 500 and DAX daily returns from January 1985 and January 1988 to December 2022, respectively.

We consider daily returns in Table 5.1 for interpretability, whereas all estimation is performed on log-losses. Daily returns of the S&P 500 and DAX average close to zero with a standard deviation of c. 1.2% and 1.4%, respectively. The highest S&P 500 and DAX daily return of c. 11.6% and 11.4% occurred on 19.10.1987 and 16.10.1989, respectively, while the worst S&P 500 and DAX daily loss (c. -20.5% and -13.1%) materialized during the Global Financial Crises (GFC) on 13.10.2008 for both indices. The negative Skewness of both indices indicates a fat left tail, i.e. many small gains and select extreme losses. Lastly, Kurtosis of c. 19.9 and 6.4 suggests a leptokurtic distribution of both indices, common for financial return series and greater than the expected kurtosis of a normally distributed series (kurtosis = 3).

Index	Mean	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
S&P 500	0.0004	-0.2047	0.1158	0.0116	-0.7719	19.8965
DAX	0.0004	-0.1314	0.1140	0.0140	-0.1113	6.4418

Table 5.1: Statistics of S&P 500 and DAX daily returns (Jan. 1985 and 1988 - Dec. 2022)

Rolling RVaR estimation of S&P 500 log-losses with Empirical, Brazauskas et al.'s and FHS estimator (Chapter 4) confirms our theoretical findings. As both Empirical and Brazauskas et al.'s estimator are based on empirical quantiles, both demonstrate similar behaviour and RVaR estimates (re-confirmed in Section 5.2.2). Brazauskas et al.'s estimator produces slightly lower rolling RVaR estimates compared to Empirical estimator. Both estimators display a fairly smooth rolling estimation, especially for large window size (e.g. = 1 000).

The FHS estimator produces a more volatile rolling RVaR estimation, spiking for periods of market turmoil (Figure 5.4). FHS estimates are larger compared to the two other estimators most of the time (re-confirmed in Section 5.2.2).

A rolling estimation window size of 250 leads to more volatility in the estimates and wider 99% confidence bands (Appendix A.6) compared to window size of 500 or 1000. Window size of 250 appears reasonable for the Empirical and Brazauskas et al.'s estimator, but the FHS estimator seems prone to artifacts (Figure 5.1, e.g. 1988, 1992, 1994 or 2020).

Unsurprisingly, larger rolling estimation window size of 500 leads to smoother RVaR estimates for all estimators and more narrow 99% confidence bands. However, FHS estimates still seem to overestimate RVaR for some points (Figure 5.2, e.g. 1988, 1994, 2011, 2020).

We also observe smaller differences between Empirical and Brazauskas et al.'s estimator and more stable rolling FHS estimates for window size = 1 000 (Figure 5.3).

Figure 5.1: Rolling RVaR estimation of S&P 500 log-losses with EMP, BRZ and FHS estimator (window size $= 250$, $(p,q) = (0.95, 0.975)$)

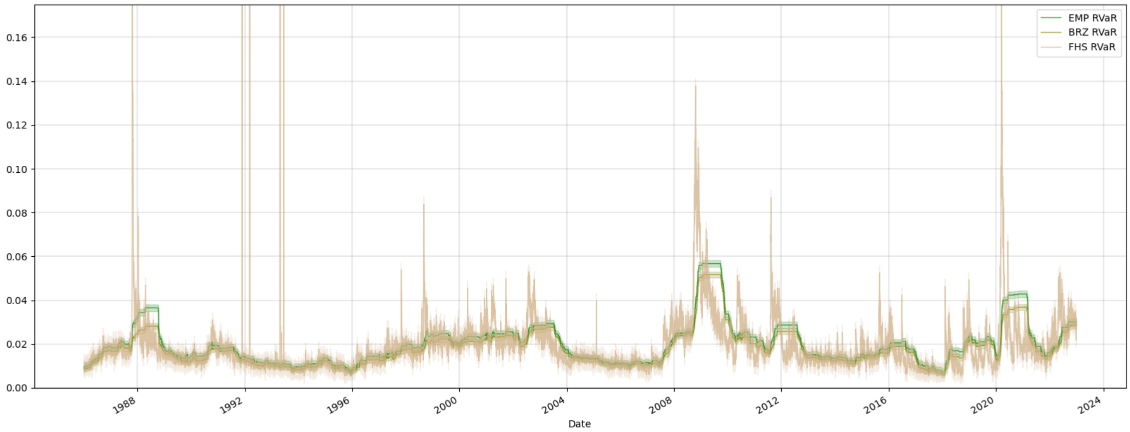


Figure 5.2: Rolling RVaR estimation of S&P 500 log-losses with EMP, BRZ and FHS estimator (window size $= 500$, $(p,q) = (0.95, 0.975)$)

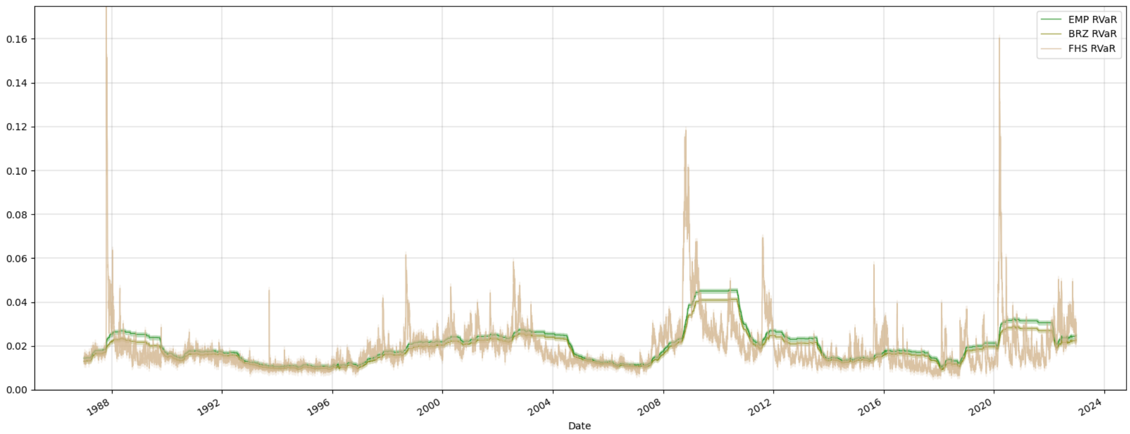
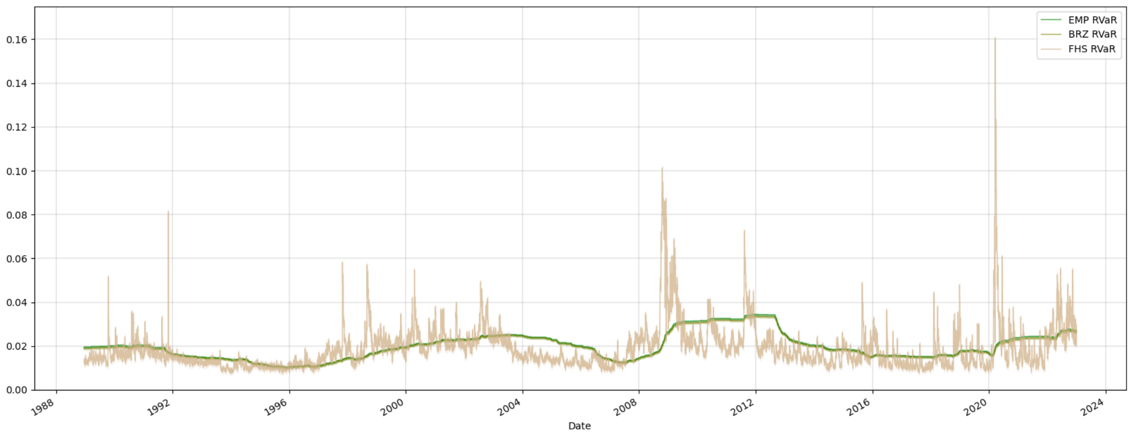
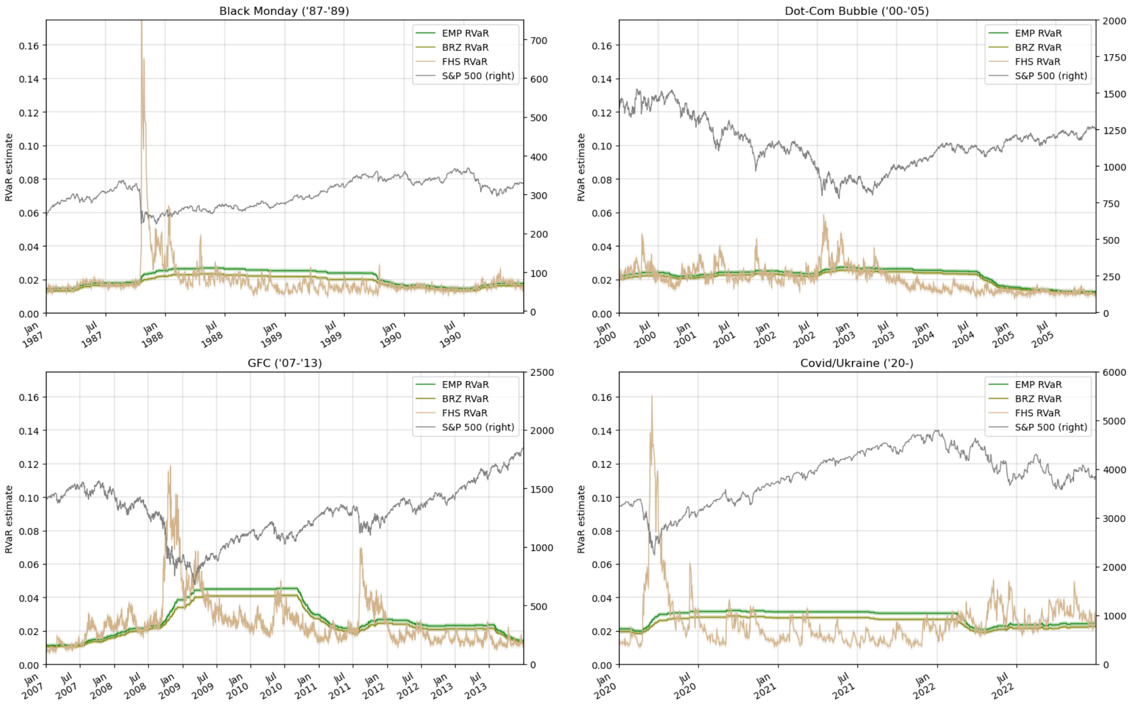


Figure 5.3: Rolling RVaR estimation of S&P 500 log-losses with EMP, BRZ and FHS estimator (window size $= 1000$, $(p,q) = (0.95, 0.975)$)



We examine the estimators during market downturns more closely in Figure 5.4, using Black-Monday, Dot-Com Bubble, GFC and Covid/Ukraine crises as examples for periods of significant market movements and hence, volatile RVaR (S&P 500 price level included as grey line). All three estimators reflect severe market downturns in the RVaR estimate. The more volatile FHS estimator tends to spike for the periods of biggest losses but returns to 'baseline' fairly quickly, even if the index has not recovered. This behaviour can be explained by the one-step volatility forecast we apply (Section 4.3). Both Empirical and Brazauskas et al.'s estimator stay elevated throughout the recovery period, sometimes even beyond (e.g. Covid/Ukraine crises, while S&P 500 recovers by mid-2020, RVaR estimates remain elevated until early-2022).
For illustration of the underlying ES estimates by Empirical, Brazauskas et al.'s and FHS estimator refer to Figure A.1 in the Appendix.

Figure 5.4: Rolling RVaR estimation of S&P 500 log-losses with EMP, BRZ, and FHS estimator during volatile markets (window size = 500, $(p,q) = (0.95, 0.975)$)



The results of rolling RVaR estimation of DAX log-losses confirm our theoretical findings as well and are in line with the S&P 500 results described above. As with S&P 500 estimates, Empirical and Brazauskas et al.'s estimator demonstrate similar behaviour and RVaR estimates for the DAX index. Also as above, Brazauskas et al.'s RVaR estimates are slightly lower compared to Empirical estimator. Both Empirical and Brazauskas et al.'s estimators display a fairly smooth rolling estimation, while FHS estimates are more volatile for the DAX index, spiking for periods of market turmoil (Figure 5.5).

As with S&P 500 results, smaller estimation window sizes lead to more volatility in the DAX estimates compared to larger window sizes and the FHS estimator seems to overestimate RVaR for some days, for window sizes 250 and 500 especially. Refer to the Appendix for rolling RVaR estimation for window size of 250 (Figure A.2) and 1 000 (Figure A.3).

Figure 5.5: Rolling RVaR estimation of DAX log-losses with EMP, BRZ and FHS estimator (window size = 500, $(p,q) = (0.95, 0.975)$)

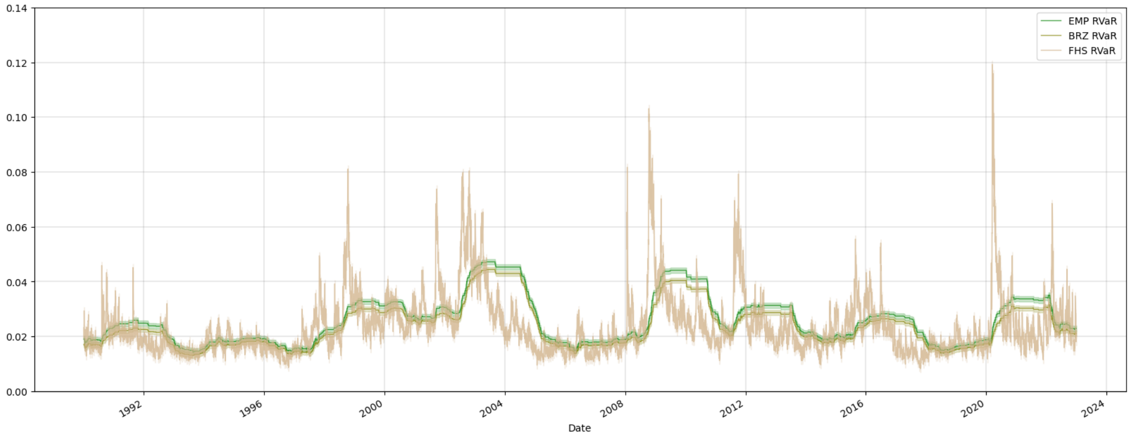
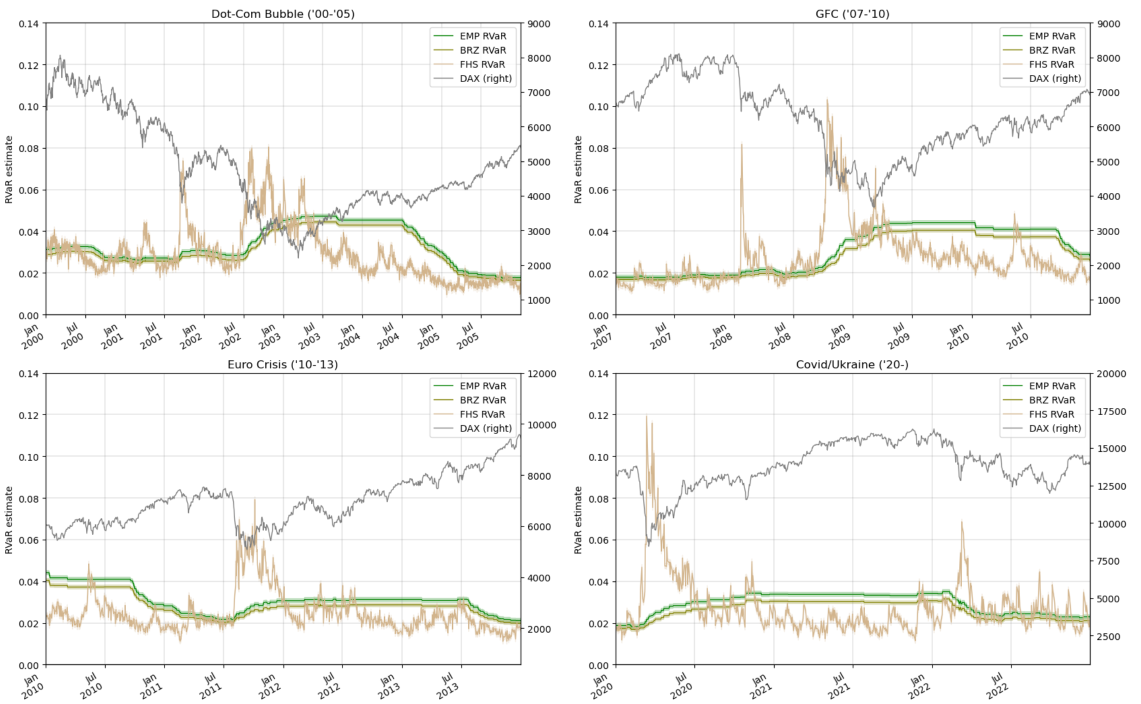


Figure 5.6: Rolling RVaR estimation of DAX log-losses with EMP, BRZ, and FHS estimator during volatile markets (window size = 500, $(p,q) = (0.95, 0.975)$)



Again, we examine the estimators during market downturns (Figure 5.6) using Dot-Com Bubble, GFC, Euro Crisis and Covid/Ukraine crises as examples for periods of volatile RVaR (DAX index level included as grey line). As with S&P 500, all estimators reflect severe market downturns in the RVaR estimate. The more volatile FHS estimates tend to spike for the periods of biggest losses but return to 'baseline' once the downward trend is broken, even if the index has not yet recovered. Again, this behaviour can be explained by the one-step volatility forecast we apply (Section 4.3). Both Empirical and Brazauskas et al.'s estimator stay elevated throughout the recovery period, sometimes even beyond (e.g. Covid/Ukraine crises, while DAX recovers by 2021, RVaR estimates remain elevated until early-2022).

5.2.2 Numerical results

This subsection presents the Monte Carlo simulation study results. We fit model parameters via Maximum-Likelihood estimation on the S&P 500 log-loss data, yielding $\mathcal{N}(\mu = -0.000328, \sigma^2 = 0.011653)$, $t(\nu = 15101953)$ and a GARCH $(1,1)$ with constant mean and Gaussian residuals, in particular $\mu = 0.100700$, $\alpha_0 = 0.000003$, $\alpha_1 = 0.099999$ and $\beta_1 = 0.879987$ (Appendix A.10).

No estimator uniformly outperforms across scenarios. Brazauskas et al.'s and FHS estimator seem best for Gaussian settings, FHS estimator for Student's-t scenarios, while Empirical and Brazauskas et al.'s estimates seem most accurate for GARCH$(1,1)$ settings. In most scenarios at least one estimator predicts RVaR with good accuracy (Relative Bias $< 1\%$). Except FHS estimator in GARCH$(1,1)$ scenarios, all estimators predict RVaR with decent accuracy (Relative Bias $< 10\%$). All estimators seem to perform better with more data within RVaR threshold, i.e. for larger n and larger delta of α and β. As with the empirical study, we observe similar outcomes with Empirical and Brazauskas et. al.'s estimator. Same reasoning as in section 5.2.1 applies. While Empirical and Brazauskas et al.'s estimator over- and underestimate true RVaR across scenarios, FHS estimator tends to overestimate RVaR (negative Bias) for Gaussian and Students-t scenarios and underestimates RVaR in GARCH$(1,1)$ scenarios. The variance of Brazauskas et al.'s estimator is lower than the Empirical estimator's variance in most cases, which corroborates our theoretical finding that Brazauskas et al.'s estimator is potentially more robust than the Empirical estimator (Section 4.2).

Across most Gaussian scenarios, Brazauskas et al.'s and FHS estimators perform best . Accuracy is fairly similar across all three estimators, except Brazauskas et al.'s estimator outperforms for $n = 1\,000$ for normal distribution with Relative Bias of $< 1\%$. FHS estimates demonstrate lowest RMSE across all Gaussian scenarios.

For Student's-t distribution scenarios, FHS estimator perform best for almost all significance levels and sample sizes. Again, FHS estimator displays lowest RMSE across all Student's-t scenarios. The difference in accuracy is small across the estimators, except for $n = 1\,000$, where Empirical estimator falls behind Brazauskas et al.'s and FHS estimator accuracy.

A different pattern manifests for GARCH$(1,1)$ scenarios, for which either Empirical or Brazauskas et al.'s estimator is best, depending on significance levels and sample size. FHS estimator underestimates RVaR across GARCH$(1,1)$ scenarios significantly compared to the other two estimators.

Based on this numerical study, we recommend the FHS estimator for RVaR estimation in a Gaussian and Student's-t setting and recommend Empirical and Brazauskas et al.'s estimator for GARCH scenarios.

The following tables show Bias, Relative Bias, RMSE and Variance (Appendix A.3), grouped by distribution, sample size and significance level. Best results per scenario are bold.

Table 5.2: Results for distribution $\mathcal{N}(\mu,\sigma)$, $n = 250$.

(α,β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	**-0.000029**	**-0.001401**	0.001602	0.000002
$(5.0\%, 2.5\%)$	BRZ	0.001254	0.061139	0.001977	0.000002
	FHS	-0.000353	-0.017233	**0.001447**	0.000002
	EMP	0.000490	0.022316	0.001769	0.000003
$(5.0\%, 1.0\%)$	BRZ	0.000520	0.023697	0.001789	0.000003
	FHS	**-0.000132**	**-0.005997**	**0.001503**	0.000002
	EMP	0.001250	0.051277	0.002347	0.000004
$(2.5\%, 1.0\%)$	BRZ	-0.000827	-0.033943	0.002278	0.000005
	FHS	**0.000123**	**0.005049**	**0.001892**	0.000004

Table 5.3: Results for distribution $\mathcal{N}(\mu,\sigma)$, $n = 500$.

(α,β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	**-0.000415**	**-0.020240**	0.001219	0.000001
$(5.0\%, 2.5\%)$	BRZ	0.000880	0.042908	0.001407	0.000001
	FHS	-0.000523	-0.025496	**0.001143**	0.000001
	EMP	-0.000251	-0.011419	0.001270	0.000001
$(5.0\%, 1.0\%)$	**BRZ**	**0.000103**	**0.004680**	0.001232	0.000001
	FHS	-0.000259	-0.011779	**0.001119**	0.000001
	EMP	-0.000322	-0.013226	0.001535	0.000002
$(2.5\%, 1.0\%)$	BRZ	-0.001552	-0.063672	0.002209	0.000003
	FHS	**-0.000102**	**-0.004202**	**0.001382**	0.000002

Table 5.4: Results for distribution $\mathcal{N}(\mu, \sigma)$, $n = 1\,000$.

(α, β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	-0.000185	-0.008999	0.000843	0.000001
$(5.0\%, 2.5\%)$	**BRZ**	**0.000049**	**0.002378**	0.000820	0.000001
	FHS	-0.000268	-0.013047	**0.000774**	0.000001
	EMP	-0.000251	-0.011422	0.000891	0.000001
$(5.0\%, 1.0\%)$	**BRZ**	**0.000015**	**0.000668**	0.000847	0.000001
	FHS	-0.000304	-0.013840	**0.000804**	0.000001
	EMP	-0.000291	-0.011945	0.001088	0.000001
$(2.5\%, 1.0\%)$	**BRZ**	**0.000030**	**0.001249**	0.001029	0.000001
	FHS	-0.000327	-0.013438	**0.001011**	0.000001

Table 5.5: Results for distribution $t(\nu)$, $n = 250$.

(α, β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	-0.006760	-0.003782	0.136945	0.018061
$(5.0\%, 2.5\%)$	BRZ	0.103440	0.057865	0.167781	0.017102
	FHS	**-0.005562**	**-0.003111**	**0.121080**	0.014086
	EMP	0.038612	0.020194	0.151782	0.020034
$(5.0\%, 1.0\%)$	BRZ	0.040475	0.021168	0.150872	0.020170
	FHS	**0.012808**	**0.006698**	**0.129834**	0.016412
	EMP	0.100731	0.047525	0.192549	0.026966
$(2.5\%, 1.0\%)$	BRZ	-0.081836	-0.038611	0.193187	0.031883
	FHS	**0.034318**	**0.016191**	**0.159228**	0.024228

Table 5.6: Results for distribution $t(\nu)$, $n = 500$.

(α, β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	-0.028623	-0.016012	0.104393	0.009081
$(5.0\%, 2.5\%)$	BRZ	0.082503	0.046152	0.126552	0.008205
	FHS	**-0.013304**	**-0.007442**	**0.091128**	0.007179
	EMP	-0.029054	-0.015195	0.107101	0.010865
$(5.0\%, 1.0\%)$	BRZ	0.000990	0.000518	0.102408	0.010565
	FHS	**-0.000004**	**-0.000002**	**0.091545**	0.008745
	EMP	**-0.013251**	**-0.006252**	0.130985	0.016569
$(2.5\%, 1.0\%)$	BRZ	-0.119825	-0.056534	0.180645	0.018321
	FHS	0.033694	0.015897	**0.122551**	0.013900

Table 5.7: Results for distribution $t(\nu)$, $n = 1\,000$.

(α, β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	-0.020317	-0.011366	0.072112	0.004955
$(5.0\%, 2.5\%)$	BRZ	-0.000428	-0.000239	0.068939	0.004908
	FHS	**-0.000198**	**-0.000111**	**0.061629**	0.004075
	EMP	-0.021480	-0.011234	0.076852	0.005737
$(5.0\%, 1.0\%)$	BRZ	0.000821	0.000430	0.073266	0.005602
	FHS	**-0.000134**	**-0.000070**	**0.065846**	0.004467
	EMP	-0.027653	-0.013047	0.094421	0.008672
$(2.5\%, 1.0\%)$	BRZ	-0.000825	-0.000389	0.089107	0.008439
	FHS	**-0.000507**	**-0.000239**	**0.084236**	0.007367

Table 5.8: Results for distribution GARCH$(1,1)$, $n = 250$.

(α,β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	**0.002681**	**0.022671**	**0.003611**	0.000033
$(5.0\%, 2.5\%)$	BRZ	0.004600	0.038484	0.005020	0.000027
	FHS	0.099925	0.830283	0.099947	0.000031
	EMP	0.004545	0.037975	0.005483	0.000038
$(5.0\%, 1.0\%)$	**BRZ**	**-0.000578**	**-0.004187**	**0.002865**	0.000036
	FHS	0.099765	0.820438	0.099802	0.000036
	EMP	**0.007816**	**0.063697**	**0.008612**	0.000043
$(2.5\%, 1.0\%)$	BRZ	-0.009005	-0.072140	0.009768	0.000048
	FHS	0.099616	0.805543	0.099673	0.000042

Table 5.9: Results for distribution GARCH$(1,1)$, $n = 500$.

(α,β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	**-0.000345**	**-0.002462**	**0.002631**	0.000026
$(5.0\%, 2.5\%)$	BRZ	0.004078	0.034187	0.004606	0.000021
	FHS	0.099615	0.827844	0.099641	0.000023
	EMP	**-0.000517**	**-0.003861**	**0.002672**	0.000030
$(5.0\%, 1.0\%)$	BRZ	-0.001264	-0.010040	0.002692	0.000028
	FHS	0.099418	0.817221	0.099444	0.000027
	EMP	**-0.001006**	**-0.007556**	**0.003801**	0.000051
$(2.5\%, 1.0\%)$	BRZ	-0.010342	-0.082853	0.011007	0.000052
	FHS	0.098995	0.799146	0.099047	0.000044

Table 5.10: Results for distribution GARCH$(1,1)$, $n = 1\,000$.

(α, β)	**Estimator**	**Bias**	**R. Bias**	**RMSE**	σ^2
	EMP	-0.001026	-0.008375	0.001780	0.000011
$(5.0\%, 2.5\%)$	**BRZ**	**-0.000857**	**-0.006987**	**0.001623**	0.000010
	FHS	0.099829	0.829461	0.099837	0.000010
	EMP	-0.001460	-0.011801	0.002338	0.000016
$(5.0\%, 1.0\%)$	**BRZ**	**-0.001415**	**-0.011447**	**0.002208**	0.000015
	FHS	0.099283	0.815697	0.099296	0.000015
	EMP	**-0.002270**	**-0.017974**	0.003758	0.000026
$(2.5\%, 1.0\%)$	BRZ	-0.002426	-0.019256	**0.003686**	0.000023
	FHS	0.098277	0.792555	0.098315	0.000023

6 Conclusion

In the discussion around which risk measure is best, RVaR poses an interesting alternative as a robust risk measure that is sensitive to potential losses in the tail and contains VaR and ES as special cases. While RVaR is relevant for researchers and practitioners, peer-reviewed publications on estimating RVaR are sparse. The thesis expands on Biswas & Sen (2023), Brazauskas et al. (2008), Chen (2007), Giannopoulos & Tunaru (2005) a.o., comparing three non-parametric RVaR estimators' theoretical properties and behaviour in application via empirical and MC simulation study.

We establish point-wise consistency and asymptotic normality of the Empirical estimator, assuming absolute continuity of the loss distribution. Further, we show that Brazauskas et al.'s estimator is point-wise consistent and asymptotically normal as well, however, absolute continuity of the loss distribution is not assumed. Thus, Brazauskas et al.'s estimator is more robust under partially discontinuous loss distributions compared to the Empirical Estimator, which our numerical study corroborates.
We analyse the FHS estimator representing a numerical method in contrast to the two closed form estimators. Although FHS estimator is not proven to be consistent or asymptotically normal for RVaR, FHS estimates outperform in most Gaussian and Student's-t settings in our numerical study. For losses following a GARCH$(1,1)$ model however, Empirical and Brazauskas et al.'s estimator are most accurate in our MC study.
Our empirical study shows that FHS produces more volatile rolling estimates compared to fairly smooth Empirical and Brazauskas et al.'s estimates and is more computationally intensive.

Future studies can apply the RVaR backtesting framework recently proposed in Biswas & Sen (2023, amended Feb. '23) to Empirical and Brazauskas et al.'s estimator and test different conditional volatility models or more elaborate volatility forecasting within FHS method. Furthermore of interest is testing more distribution scenarios in numerical studies as Biswas & Sen (2023) suggest that e.g., FHS estimator struggles with heavy-tailed distributions. Lastly, we suggest exploring theoretical aspects of RVaR estimation for which peer-reviewed research is missing, such as closed form formulas for estimators' variance, estimator behaviour for small samples and backtesting frameworks for RVaR.

A Appendix

A.1 On risk measures

Definition A.1 (Lévy distance, (Cont et al. 2010))**.** The Lévy distance between two CDFs $F, G \in \mathcal{D}$ is

$$d(F,G) := \inf\{\varepsilon > 0 \mid \forall x \in \mathbb{R}:\ F(x-\varepsilon) - \varepsilon \leq G(x) \leq F(x+\varepsilon) + \varepsilon\}.$$

Proof: VaR is monotone, translation invariant and positive homogeneous but not subadditive. Let $X, Y \in \mathcal{F}$, $a, b \in \mathbb{R}$ and $b > 0$. VaR satisfies monotonicity, as $X \leq Y$ implies

$$\mathbb{P}(Y \leq k) = \mathbb{P}(X \leq Y \leq k) \leq \mathbb{P}(X \leq k), \quad \text{for } k \geq 0,$$

therefore,

$$\mathbb{P}(Y \leq k) \geq p \Rightarrow \mathbb{P}(X \leq k) \geq p, \quad \text{for } k \geq 0,$$

which proves $\mathrm{VaR}_p(X) \leq \mathrm{VaR}_p(Y)$ as VaR is minimal. Translation invariance and positive homogeneity are satisfied as,

$$\begin{aligned} \mathrm{VaR}_p(a+bX) &= \inf\{x \mid \mathbb{P}(a+bX \leq x) \geq p\} = \inf\{x \mid \mathbb{P}(X \leq (x-a)/b) \geq p\} \\ &= \inf\{a+bs \mid \mathbb{P}(X \leq s) \geq p\} = a + b \cdot \inf\{s \mid \mathbb{P}(X \leq s) \geq p\} \\ &= a + b \cdot \mathrm{VaR}_p(X). \end{aligned}$$

Thus VaR is a monetary risk measure. To illustrate VaR's lack of subadditivity, we quote the example given in Yan (2018, p.265): Let X and Y be two i.i.d. random variables satisfying

$$\mathbb{P}(X=0) = 0.95,\ \mathbb{P}(X=1) = 0.05,$$

then

$$\mathrm{VaR}_{0.95}(X) = \mathrm{VaR}_{0.95}(Y) = 0.$$

However, since

$$\mathbb{P}(X+Y=0) = 0.9025,\ \mathbb{P}(X+Y=1) = 0.095,\ \mathbb{P}(X+Y=2) = 0.0025,$$

we have $\mathrm{VaR}_{0.95}(X+Y) = 1 \not\leq 0 = \mathrm{VaR}_{0.95}(X) + \mathrm{VaR}_{0.95}(Y)$. □

Proof: Statement (3.3). Closely following McNeil (2015, Lemma 2.13), let U denote a random variable with uniform distribution on $[0,1]$. We use $\mathbb{E}[X;A] := \mathbb{E}[XI_A]$ for any generic set $A \in \mathcal{F}$. It is a well-known fact from elementary probability theory that the random variable $F_X^{-1}(U)$ has CDF F_X (Lemma A.1). We first show that $(1-p)\mathrm{ES}_p(X) = \mathbb{E}[X; X \geq \mathrm{VaR}_p(X)]$. Now,

$$\mathbb{E}[X; X \geq \mathrm{VaR}_p(X)] = \mathbb{E}\left[F_X^{-1}(U); F_X^{-1}(U) \geq F_X^{-1}(p)\right] = \mathbb{E}\left[F_X^{-1}(U); U \geq p\right].$$

In the last equality we use the fact that F_X^{-1} is strictly increasing since F_X is continuous. Thus we get $\mathbb{E}\left[F_X^{-1}(U); U \geq p\right] = \int_p^1 F_X^{-1}(u)\, du$. Lastly,

$$\begin{aligned} \mathrm{ES}_p(X) &= \frac{1}{1-p}\mathbb{E}[X; X \geq \mathrm{VaR}_p(X)] = \frac{1}{\mathbb{P}\left(X \geq F_X^{-1}(p)\right)}\mathbb{E}[X; X \geq \mathrm{VaR}_p(X)] \\ &= \mathbb{E}[X \mid X \geq \mathrm{VaR}_p(X)], \end{aligned}$$

by definition of conditional expectation and density. Note that we assume continuity of F_X. A proof that (3.3) does not hold when F_X is not continuous can be found in Brazauskas et al. (2008, Appenidx A). □

Proof: ES is a coherent risk measure. Monotonicity of ES follows directly by monotonicity of VaR. Translation invariance and positive homogeneity are satisfied as VaR enjoys these properties and integrals are linear. To prove subadditivity of ES, we recount one of the arguments given in Embrechts & Wang (2015), which requires Lemma A.1, a classic result in probability theory.

Lemma A.1 (Proposition 7.2 in McNeil (2015))**.** *For any random variable X, there exists a $U[0,1]$ random variable U_X such that $X = F_X^{-1}(U_X)$ almost surely.*

Following, denote $A_X = \mathbf{I}_{\{U_X \geq p\}} \in Ber(1-p)$. We now show that

$$\mathbb{E}[XA_X] \geq \mathbb{E}[XB] \quad \text{for all } B \in Ber(1-p). \tag{A.1}$$

Since $\mathbb{E}[A_X - B] = 0$, we have $\mathbb{E}[X(A_X - B)] = \mathbb{E}[(X-m)(A_X - B)]$ for all $m \in \mathbb{R}$. Take $m = F_X^{-1}(p)$. Utilizing Lemma A.1, if $F_X^{-1}(U_X) > m$, then $U_X > p$, $A_X = 1$ and $\mathbb{E}[(X-m)(A_X-B)] \geq 0$; if $F_X^{-1}(U_X) < m$, then $U_X < p$, $A_X = 0$ and $\mathbb{E}[(X-m)(A_X-B)] \geq 0$; if $F_X^{-1}(U_X) = m$, then $\mathbb{E}[(X-m)(A_X-B)] = 0$. In summary, $\mathbb{E}[X(A_X-B)] = \mathbb{E}[(X-m)(A_X-B)] \geq 0$ which concludes statement (A.1).
Further, we have

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 F_X^{-1}(u)\,du = \frac{1}{1-p}\mathbb{E}\left[F_X^{-1}(U_X) I_{\{U_X \geq p\}}\right] = \frac{1}{1-p}\mathbb{E}[XA_X].$$

From equation (A.1),

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\sup\left\{\mathbb{E}[XB] : B \in Ber(1-p)\right\},\ X \in L^\infty.$$

That is, ES_p is the supremum of the additive maps $X \mapsto \frac{1}{1-p}\mathbb{E}[XB]$ over $B \in Ber(1-p)$, and hence is subadditive. □

Proof: RVaR properties. RVaR is monotone, translation invariant and positive homogeneous by linearity and monotonicity of integrals and VaR's properties. As VaR, RVaR is not subadditive. □

A.2 On estimators for Range Value at Risk

A.2.1 Empirical estimator

We define the *empirical distribution function*, closely following van der Vaart (2012, Ch. 19.1). Let $X_1, \ldots, X_n$ be a random sample of observed losses from a CDF $F_X \in \mathcal{D}$. The *empirical distribution function* is defined as

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}_{\{X_i \leq x\}}. \tag{A.2}$$

If the underlying distribution F is completely unknown, this is the natural estimator. This estimator is unbiased as $n\hat{F}(x)$ follows a binomial distribution with mean $nF_X(x)$,

$$\mathbb{E}\left[n\hat{F}(x)\right] = \mathbb{E}\left[\sum_{i=1}^{n} \mathbf{I}_{\{X_i \leq x\}}\right] = \sum_{i=1}^{n} \mathbb{P}(X_i \leq x)$$
$$= \sum_{i=1}^{n} F_X(x) = nF_X(x). \tag{A.3}$$

Further, this estimator is point-wise consistent, as applying the law of large numbers yields

$$\forall x: \quad \hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}_{\{X_i \leq x\}} \xrightarrow[\text{a.s.}]{n\to\infty} \mathbb{E}\left[\mathbf{I}_{\{X_i \leq x\}}\right]$$
$$= \mathbb{P}(X_i \leq x) = F_X(x). \tag{A.4}$$

Theorem A.1 (Glivenko-Cantelli theorem; Theorem 19.1 van der Vaart (2012))**.** *If $X_1, X_2, \ldots$ are i.i.d. random variables with distribution function F, then $\|\hat{F} - F\|_\infty \xrightarrow{a.s.} 0$ with $\hat{F}$ as in (A.2) and $\|\circ\|_\infty$ the supremum norm.*

Lemma A.2 (Lemma 21.1 in van der Vaart (2012))**.** *For every $0 < p < 1$ and $x \in \mathbb{R}$,*

(i) $F^{-1}(p) \leq x$ iff $p \leq F(x)$;

(ii) $F \circ F^{-1}(p) \geq p$ with equality iff p is in the range of F; equality can fail only if F is discontinuous at $F^{-1}(p)$;

(iii) $F\_ \circ F^{-1}(p) \leq p$;

(iv) $F^{-1} \circ (x) \leq x$; equality fails iff x is in the interior or at the right end of a 'flat' of F;

(v) $F^{-1} \circ F \circ F^{-1} = F^{-1}$; $F \circ F^{-1} \circ F = F$;

(vi) $(F \circ G)^{-1} = G^{-1} \circ F^{-1}$.

Lemma A.3 (Delta method, Theorem 3.1 van der Vaart (2012))**.** *Let $\phi : \mathbb{W} \subset \mathbb{R}^k \mapsto \mathbb{R}^m$ be a map defined on a subset of $\mathbb{R}^k$ and differentiable at θ. Let T_n be a sequence of random vectors taking their values in the domain of ϕ. If $r_n(T_n - \theta) \xrightarrow{d} T$ for numbers $r_n \to \infty$, then $r_n(\phi(T_n) - \phi(\theta)) \xrightarrow{d} \phi'(T)$. In particular,*

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right) \quad \Rightarrow \quad \sqrt{n}(\phi(T_n) - \phi(\theta)) \xrightarrow{d} \mathcal{N}\left(0, [\phi'(\theta)]^2 \cdot \sigma^2\right).$$

A.2.2 Brazauskas et al.'s estimator

Lemma A.4 (Characterization of weak convergence, p.64-65 Shorack & Wellner (2009))**.** *Let $\mathfrak{F}_1 := \left\{F | F \text{ is a CDF such that} \int_{-\infty}^{\infty} |x| dF(x) < \infty\right\}$ and $F, F_1, F_2, \ldots \in \mathfrak{F}_1$. The following holds,*

$$\int_0^1 |F_n^{-1}(t) - F^{-1}(t)|\, dt \quad \xrightarrow[n\to\infty]{a.s.} \quad 0 \quad \textit{if and only if}$$
$$F_n \xrightarrow[n\to\infty]{d} F \quad \textit{and} \quad \int |x|\, dF_n \xrightarrow[n\to\infty]{a.s.} \int |x|\, dF.$$

Lemma A.5 (Section 3 in Csörgő & Zitikis (1996))**.** *The Vervaat process is defined for* $0 \leq t \leq 1$ *as*

$$V_n(t) := \int_0^t \left(\hat{F}^{-1}(s) - F^{-1}(s)\right) ds + \int_0^{F^{-1}(t)} \left(\hat{F}(x) - F(x)\right) ds.$$

It holds that

$$0 \leq V_n(t) = \int_{\hat{F}^{-1}(t)}^{F^{-1}(t)} \hat{F}(x) - t \, dx \leq |\hat{F}(F^{-1}(t) - t| \cdot |\hat{F}^{-1}(t) - F^{-1}(t)|.$$

A.3 On empirical and simulation study

Lemma A.6 (Confidence bands)**.** *Let* $\hat{\rho}_{p,q}$ *be an estimator for* $RVaR_{p,q}$ *that is asymptotically normal. The* $(1-\delta) \times 100\%$ *level asymptotic confidence interval for* $RVaR_{p,q}$ *is given by:*

$$\hat{\rho}_{p,q} \pm \frac{z_{\delta/2} \sigma_X(p,q)}{\sqrt{n}}, \tag{A.5}$$

where $z_{\delta/2}$ *is the* $(1-\delta/2) \times 100\%$ *percentile of a standard normal distribution and* n *the rolling estimation window size.*
As the standard deviation $\sigma_X(p,q)$ *is unknown, we estimate it with the observed standard deviation of the estimator as closed form variance formulas for the estimators outlined in section 4 are not yet researched to our knowledge.*

Figure A.1: Rolling ES estimation of S&P 500 log-losses with EMP, BRZ, and FHS estimator ($p = 0.95$)

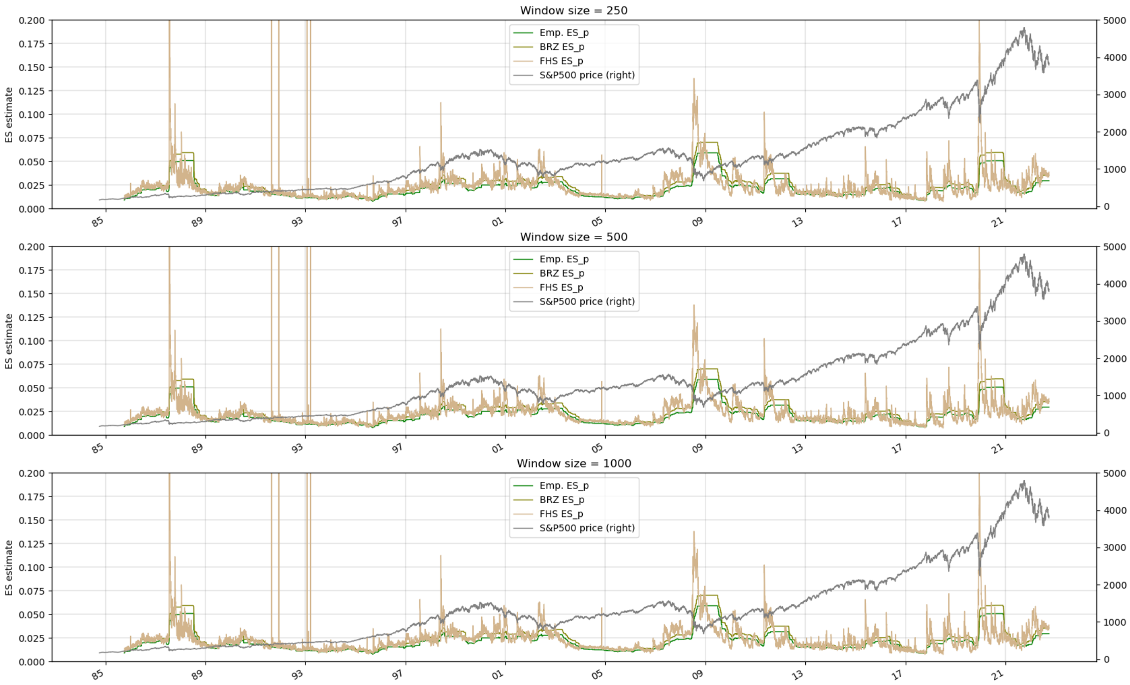


Figure A.2: Rolling RVaR estimation of DAX log-losses with EMP, BRZ and FHS estimator (window size = 250, $(p,q) = (0.95, 0.975)$)

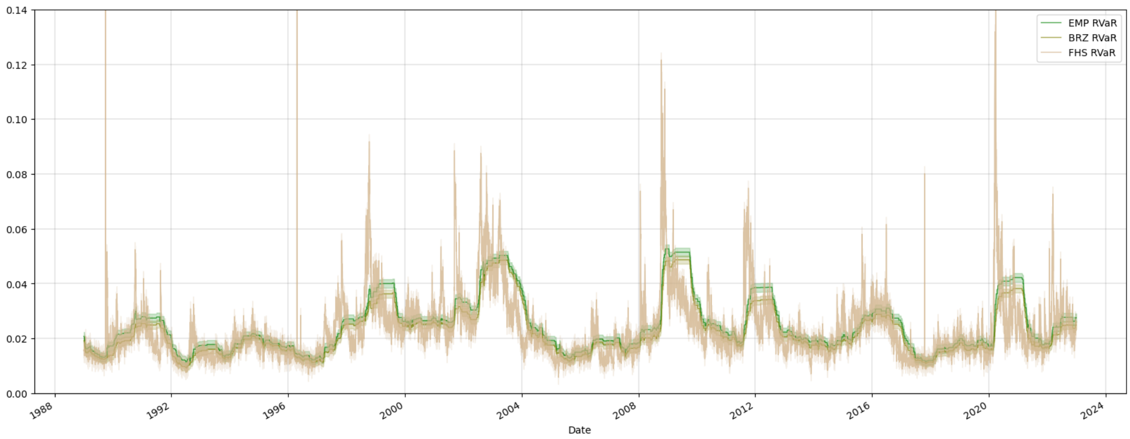
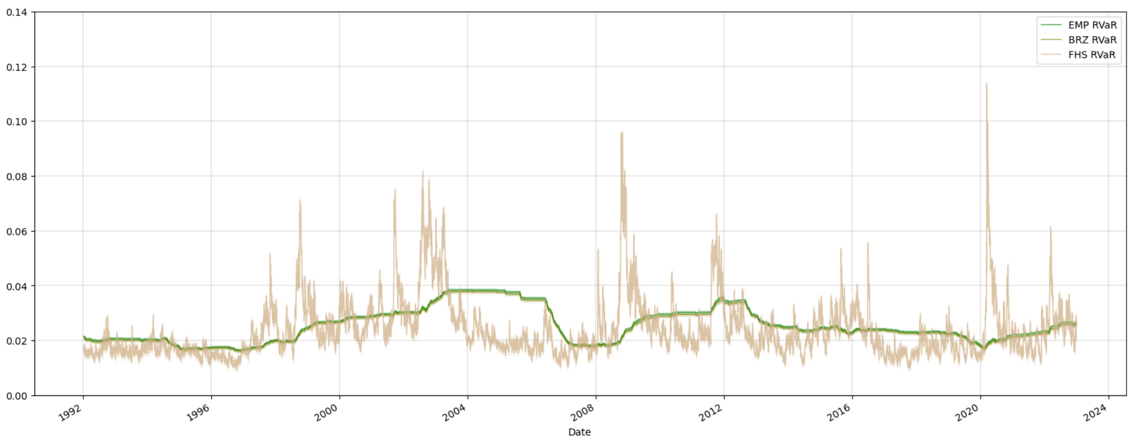


Figure A.3: Rolling RVaR estimation of DAX log-losses with EMP, BRZ and FHS estimator (window size = 1000, $(p,q) = (0.95, 0.975)$)



Definition A.2. Let $\hat{\rho}$ be an estimator used to estimate a monetary risk measure ρ. *Bias*, *Relative Bias*, *Root-Mean-Squared-Error* (RMSE) and *variance* (σ^2) are defined as:

$$\text{Bias}(\hat{\rho},\rho) = \mathbb{E}\left[\rho - \hat{\rho}\right]; \tag{A.6}$$

$$\text{R. Bias}(\hat{\rho},\rho) = \mathbb{E}\left[\frac{\rho - \hat{\rho}}{\rho}\right]; \tag{A.7}$$

$$\text{RMSE}(\hat{\rho},\rho) = \mathbb{E}\left[\sqrt{(\rho - \hat{\rho})^2}\right]; \tag{A.8}$$

$$\sigma^2(\hat{\rho}) = \mathbb{E}\left[(\hat{\rho} - \mathbb{E}[\hat{\rho}])^2\right]. \tag{A.9}$$

We use the following notation for GARCH models throughout the paper (McNeil 2015, Chapter 4).

Definition A.3 (GARCH process)**.** Let $(Z_t)_{t\in\mathbb{Z}}$ be a strict white noise process, SWN $(0,1)$. The process $(X_t)_{t\in\mathbb{Z}}$ is a GARCH(p,q) process if it is strictly stationary and satisfies for all $t \in \mathbb{Z}$ and some strictly positive-valued process $(\sigma_t)_{t\in\mathbb{Z}}$:

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1,\ldots,p$, and $\beta_j \geq 0$, $j = 1,\ldots,q$.
In particular, a GARCH$(1,1)$ with constant mean can be written as,

$$X_t = \sigma_t Z_t \tag{A.10}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{A.11}$$

Bibliography

References

Acerbi, C. & Szekely, B. (2014), 'Back-testing Expected Shortfall', *Risk* **27**(11), 76–81.

Acerbi, C. & Szekely, B. (2017), 'General properties of Backtestable Statistics', *SSRN Electronic Journal* .

Acerbi, C. & Tasche, D. (2002), 'On the coherence of expected shortfall', *Journal of Banking & Finance* **26**(7), 1487–1503.

Altman, N. & Léger, C. (1995), 'Bandwidth selection for kernel distribution function estimation', *Journal of Statistical Planning and Inference* **46**(2), 195–214.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent Measures of Risk', *Mathematical Finance* **9**(3), 203–228.

Bairakdar, R., Cao, L. & Mailhot, M. (2020), 'Range Value-at-Risk: Multivariate and Extreme Values'.

Barendse, S. (2020), 'Efficiently weighted estimation of Tail and interquantile Expectations', *Econometrics: Econometric & Statistical Methods* .

Barone-Adesi, G., Bourgoin, F. & Giannopoulos, K. (1998), 'Don't look back', *Risk* **11**, 100–103.

Barone-Adesi, G., Giannopoulos, K. & Vosper, L. (1999), 'VaR without correlations for portfolios of derivative securities', *The journal of futures markets* **5**, 583–602.

BCBS (1996), 'Amendment to the capital accord to incorporate market risks'.

BCBS (2013), 'Fundamental review of the trading book: A revised market risk framework'.

Bernard, C., Kazzi, R. & Vanduffel, S. (2020), 'Range Value-at-Risk bounds for unimodal distributions under partial information', *Insurance: Mathematics and Economics* **94**, 9–24.

Biswas, S. & Sen, R. (2019), 'Kernel based estimation of Spectral Risk Measures'.

Biswas, S. & Sen, R. (2023), 'Nonparametric Estimation of Range Value at Risk', *Computation* **11**(2), 28.

Bosq, D. (1998), *Nonparametric statistics for stochastic processes*, Springer.

Brazauskas, V., Jones, B. L., Puri, M. L. & Zitikis, R. (2008), 'Estimating conditional tail expectation with actuarial applications in view', *Journal of Statistical Planning and Inference* **138**(11), 3590–3604.

Chen, S. X. (2007), 'Nonparametric Estimation of Expected Shortfall', *Journal of Financial Econometrics* **6**(1), 87–107.

Cont, R., Deguest, R. & Scandolo, G. (2010), 'Robustness and sensitivity analysis of risk measurement procedures', *Quantitative Finance* **10**(6), 593–606.

Csörgő, M. & Zitikis, R. (1996), 'Strassen's LIL for the Lorenz Curve', *Journal of Multivariate Analysis* **59**(1), 1–12.

Davis, M. H. A. (2016), 'Verification of internal risk measure estimates', *Statistics & Risk Modeling* **33**(3-4), 67–93.

Doukhan, P. (2012), *Mixing Properties and Examples*, Springer London, Limited.

Dowd, K. (2007), *Measuring Market Risk*, Wiley & Sons, Incorporated, John.

Embrechts, P., Liu, H. & Wang, R. (2018), 'Quantile-based Risk Sharing', *Operations Research* **66**(4), 936–949.

Embrechts, P., McNeil, A. J. & Straumann, D. (2002), Correlation and dependence in Risk Management: Properties and Pitfalls, *in* 'Risk Management', Cambridge University Press, pp. 176–223.

Embrechts, P. & Wang, R. (2015), 'Seven proofs for the subadditivity of Expected Shortfall', *Dependence Modeling* **3**(1).

Emmer, S., Kratz, M. & Tasche, D. (2015), 'What is the best risk measure in practice? A comparison of standard measures', *Journal of Risk* **18**(2).

Fissler, T. & Ziegel, J. F. (2016), 'Higher order elicitability and Osband's principle', *The Annals of Statistics* **44**(4).

Fissler, T. & Ziegel, J. F. (2021), 'Evaluating Range Value at Risk forecasts', *Statistics & Risk Modeling* .

Giannopoulos, K. & Tunaru, R. (2005), 'Coherent risk measures under filtered historical simulation', *Journal of Banking & Finance* **29**(4), 979–996.

Gneiting, T. (2011), 'Making and Evaluating Point Forecasts', *Journal of the American Statistical Association* **106**(494), 746–762.

Hill, J. B. (2013), 'Expected Shortfall Estimation and Gaussian Inference for Infinite Variance Time Series', *Journal of Financial Econometrics* **13**(1), 1–44.

Huber, P. J. & Ronchetti, E. M. (2009), *Robust statistics*, Wiley.

Kou, S., Peng, X. & Heyde, C. C. (2013), 'External Risk Measures and Basel Accords', *Mathematics of Operations Research* **38**(3), 393–417.

Krätschmer, V., Schied, A. & Zähle, H. (2012), 'Qualitative and infinitesimal robustness of tail-dependent statistical functionals', *Journal of Multivariate Analysis* **103**(1), 35–47.

Krätschmer, V., Schied, A. & Zähle, H. (2014), 'Comparative and qualitative robustness for law-invariant risk measures', *Finance and Stochastics* **18**(2), 271–295.

Krätschmer, V., Schied, A. & Zähle, H. (2015), 'Quasi-Hadamard differentiability of general risk functionals and its application', *Statistics & Risk Modeling* **32**(1), 25–47.

Li, L., Shao, H., Wang, R. & Yang, J. (2018), 'Worst-Case Range Value-at-Risk with Partial Information', *SIAM J Financial Math* **9**(1), 190–218.

McNeil, A. J. (2015), *Quantitative risk management concepts, techniques and tools*, Princeton University Press.

Mehlitz, J. S. & Auer, B. R. (2020), 'A Monte Carlo evaluation of non-parametric estimators of expected shortfall', *The Journal of Risk Finance* **21**(4), 355–397.

Rockafellar, T. R. & Uryasev, S. P. (2001), 'Conditional Value-at-Risk for general loss distributions', *SSRN Electronic Journal* .

Rosenblatt, M. (1956), 'A central limit theorem and a strong mixing condition', *Proceedings of the National Academy of Sciences* **42**(1), 43–47.

Shorack, G. R. & Wellner, J. A. (2009), *Empirical processes with applications to Statistics*, Society for Industrial and Applied Mathematics.

Stephens, D. A. (2006), Asymptotic distribution of sample quantiles. Mathematical Statistics I at McGill University.

van der Vaart, A. W. (2012), *Asymptotic Statistics*, Cambridge University Press.

Vasilev, I. & Melnikov, A. (2022), RVaR hedging and market completions, *in* 'Mathematical and Statistical Methods for Actuarial Sciences and Finance', Springer International Publishing, pp. 429–434.

Wang, R. & Wei, Y. (2020), 'Risk functionals with convex level sets', *Mathematical Finance* **30**(4), 1337–1367.

Wu, Y., Yu, W., Balakrishnan, N. & Wang, X. (2021), 'Nonparametric estimation of expected shortfall via Bahadur-type representation and Berry–Esséen bounds', *Journal of Statistical Computation and Simulation* **92**(3), 544–566.

Yamai, Y. & Yoshiba, T. (2002), 'Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization', *Monetary and Economic Studies* **20**(1), 87–121.

Yan, J.-A. (2018), *Introduction to Stochastic Finance*, Springer Singapore.

Yi, Y., Feng, X. & Huang, Z. (2014), 'Estimation of extreme value-at-risk: An EVT approach for quantile GARCH model', *Economics Letters* **124**(3), 378–381.

Yokoyama, R. (1980), 'Moment bounds for stationary mixing sequences', *Zeitschrift fuer Wahrscheinlichkeitstheorie und Verwandte Gebiete* **52**(1), 45–57.

Zenti, R. & Pallotta, M. (2001), 'Risk analysis for asset managers: Historical Simulation, the Bootstrap approach and Value at Risk calculation', *SSRN Electronic Journal* .

Ziegel, J. F. (2014), 'Coherence and elicitability', *Mathematical Finance* **26**(4), 901–918.