A Note on Nonparametric Estimation of Financial Performance

Damiano Rossello *

February 14, 2014

Abstract

Evaluating the performance of an investment portfolio is a central theme in modern fund management. First, one estimates the distribution of the portfolio risk-adjusted reward. Second, the computation of a performance measure is needed to rank a given managed fund. This step is very sensitive to outliers in the dataset used for estimation and the resulting ranking can be reversed. To limit this distortion one has to check alternative performance measures in term of their sensitivity to outliers. This correspond to different theoretical foundations of the performance measures. A statistical procedure must be integrated into these foundations, driving managers and practitioners to performance attribution which is consistent with performance persistence. In this short note we show how some of the most adopted measures, namely the Sharpe ratio, the Gain-Loss ratio, the Average-Value-at-Risk ratio and its simpler version the Value-at-Risk ratio, lack qualitative robustness in the sense of the Hampel's definition.

KEY WORDS: Reward and Risk Functionals; Performance statistics; Qualitative Robustness

1 Introduction

The literature on financial performance analysis provides different tools to its quantification. In this note we restrict ourselves to those indexes represented by the ratio of a reward measure to a risk measure, as opposed to the ratio typical reward-to-variability ratios such as the

*Corresponding Author: Department of Economics and Business, University of Catania, Corso Italia, 55, 95129 Catania, Italy, email: rossello@unict.it

Sharpe ratio. To implement the performance measurement one needs to estimate both the statistics used in numerator and the denominator through observed data. Unfortunately performance attribution based on fund ranking is affected by outliers and a bad fit to the bulk of the data can change it drastically causing undue acknowledgement of managerial skill with consequences on a macro/micro-economic level. The control of estimation errors is a fundamental issue.

Reward-to-variability risk measures such as the Sharpe ratio can be (under)overestimated making it difficult to judge the true ability of a fund manager because of high variability of the delivered ranking, see for example [21]. This leads to the notion of robustness related to separately the reward and the risk measure, besides to that of their ratio.

To illustrate intuitively the problem of data contamination when performance evaluation is involved, we give a simple numerical example in Section 2. We examine the bad behavior of four performance measures: the Sharpe ratio, the Gain-Loss ratio, the Average Value-at-Risk ratio and the analogous Value-at-Risk ratio. We only consider nonparametric estimation errors. Section 3.1 contains preliminary definitions, while Section 3.2 links the financial definition and the statistical definition of the above ratios. Section 4 introduces the classical notion of robustness focusing on our application to performance measures. After further terminology explained in Section 4.1, in Section 4.2 we state a result against the robustness of all the proposed ratios.

2 Numerical Example

Are high values of a performance criterion good indicator of fund manager's ability? To gain intuition and to fully motivate our analysis in the rest of the paper we introduce an elementary example. Let X be the random rate of return on an investment fund over a daily horizon, with probability distribution

$$\mathsf{P}(X=-0.05)=0.15, \qquad \mathsf{P}(X=0.01)=0.45, \qquad \mathsf{P}(X=0.02)=0.40.$$

We interpret the above as a model of the managerial skill in that it quantifies the willingness to capture opportunities from private of public information not reflected by asset prices. To forecast and attribute financial performance taking into account the intrinsic risk of X one can compute the ratio of expected return to the standard deviation

$$\frac{\mathsf{E}(X)}{\text{st. dev.}(X)}=0.2122,$$

a risk-adjusted reward index. Now suppose an unexpected market movement is embedded into the above model, in such a way that the 'contaminated' random return becomes $\widetilde{X}$ with distribution

$$\mathsf{P}(\widetilde{X} = -0.50) = 0.0080, \quad \mathsf{P}(\widetilde{X} = -0.05) = 0.1488, \quad \mathsf{P}(\widetilde{X} = 0.01) = 0.4464,$$
$$\mathsf{P}(\widetilde{X} = 0.02) = 0.3968.$$

In this modified model there is a small probability of -50% daily return with forecasted performance

$$\frac{\mathsf{E}(\widetilde{X})}{\text{st. dev.}(\widetilde{X})} = 0.0189.$$

To interpret these results, we can think of how outliers might hide the managerial skill concerning a high positive performance of at least 1% daily return with 84.32 % of probability. To what extent is past performance a good indicator of future performance? Using historical data may be misleading in the presence of data contamination, especially if other funds must be evaluated. Suppose an alternative investment is under consideration and that its random rate of return Y has distribution

$$\mathsf{P}(Y = -0.024) = 0.45, \quad \mathsf{P}(Y = 0.003) = 0.15, \quad \mathsf{P}(Y = 0.02) = 0.40$$

and performance

$$\frac{\mathsf{E}(Y)}{\text{st. dev.}(Y)} = 0.1603.$$

To compare these two funds we can rank their performance through the corresponding risk-adjusted reward index. If there is not data contamination then the performance of the first fund is higher. But it is clear how outliers determine nonresistant ranking, and a small probability mass placed at $\widetilde{X} = -0.50$ may induce investors to prefer the second fund based on the chosen performance criterion. What happen if we use other criteria? Taking for example the negative of quantile $-q_\alpha(\cdot)$ at the level $\alpha = 0.01$ as a risk measure we compute:

$$\frac{\mathsf{E}(X)}{-q_\alpha(X)} = 0.1025, \quad \frac{\mathsf{E}(\widetilde{X})}{-q_\alpha(\widetilde{X})} = 0.0020, \quad \frac{\mathsf{E}(Y)}{-q_\alpha(Y)} = 0.2014.$$

It seems that a different approach to quantify riskiness removes the performance criterion's sensitivity to data contamination and preserves fund ranking. It is worth noting how $-q_\alpha(\cdot)$ behaves with respect to st. dev.$(\cdot)$ when skewness and fat-tails affect the probability law of random returns.

3 Estimation of Performance Measures

3.1 Notations and Preliminary Definitions

Let $(\Omega, \mathscr{F}, \mathsf{P})$ be a probability space on which any random variable (rv) is defined to represent the return on investments under different market scenarios measured at a final date. From the estimation perspective the main concern are functionals depending upon the distribution of X, then we denote $\mathscr{D}$ the convex set of cumulative distribution functions $F_X(x) = \mathsf{P}(X \leqslant x)$ for $x \in \mathbb{R}$. We assume $\delta_x \in \mathscr{D}$ for every $x \in \mathbb{R}$, i.e. point-mass distributions are admitted. We write $X \sim F$ if and only if $F = F_X$. The empirical distribution of a random sample $\mathbf{X} = (X_1, \ldots, X_n)$ having size $n \in \{1, 2, \ldots\}$ is $\mathbb{F}_n(\omega, x) = \frac{1}{n}\sum_{i=1}^n \mathbf{I}_{\{X_i \leqslant x\}}(\omega)$, where X_i are i.i.d. rv's with common distribution F. By the Law of Large Numbers it is a consistent estimator of the unknown $F \in \mathscr{D}$; by the Glivenko-Cantelli theorem the convergence is uniform in x with probability 1. The estimate of $\mathbb{F}_n(\omega, x)$ at a given data set $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ from historical observations or Monte Carlo simulation is $\widehat{\mathbb{F}}_n(x) = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}((-\infty, x])$, or equivalently $\frac{1}{n}\sum_{i=1}^n \mathbf{I}_{\{x_i \leqslant x\}}$, indeed a discrete uniform distribution that assigns probability $\frac{1}{n}$ to each of the sample values X_i. This implies that $\mathscr{D}$ contains all the empirical df's, which from now on will be the additional assumption. Typical financial quantities such as portfolio expected return or portfolio volatility can be expressed as a **statistical functional**, i.e. a real-valued map $\theta(F)$ over those df's for which the definition makes sense. Two fundamental examples are the mean $\theta(F) = \int x \mathrm{d}F$ corresponding to the expected value $\mathsf{E}(X) = m$ and the standard deviation $\theta(F) = \left(\int (x-m)^2 \mathrm{d}F\right)^{1/2}$ corresponding to the central second moment $v^2 = \mathsf{E}(X-m)^2$, for $X \sim F$ and provided that both m, v^2 are finite. It is worth noting that this is the case when $X \in L^p$ for $p = 1, 2$, or equivalently when F belongs to the subset $\mathscr{D}^p \subset \mathscr{D}$ of those df's having finite pth moment.

3.2 Representation of Performance Measures and Their Estimators

We assume all trades to be financed at horizon and take all returns banked to the same terminal date. Thus any $X \in \mathscr{X} \subset L^0$ represents the random rate of portfolio's return from a trading strategy. Here L^0 denotes the space of all (equivalence classes of) rv's. By a **measure of performance** we mean a map $\varphi : \mathscr{X} \to [0, \infty]$, where $\varphi(X)$ represents the value of the services potentially provided by the portfolio management industry; $\mathscr{X}$ is assumed to be a convex cone containing the constants functions, in fact a vector space. Classical performance measures arise in portfolio optimization problems and are indeed expressed as reward-risk ratios. In this paper we consider those listed below.

- When a given investor solves its portfolio selection problem in terms of expected return and variance of return, the optimized investment has the highest Sharpe ratio defined as $\frac{m}{v}$, where X is usually the excess return over a (possibly random) benchmark. The Sharpe ratio is often used for ranking alternative investments and decide what portfolio has experienced the best performance, i.e. the highest risk-adjusted return within the asset universe.

- Since the Sharpe ratio violates the no-arbitrage condition, Bernardo and Ledoit (2000) proposed the Gain-Loss ratio $\frac{m}{m_-}$ to correct this anomaly, where $m_- := \mathsf{E}(-\min\{X,0\})$ and the corresponding statistical functional is $\int x_- \mathrm{d}F$, whit $x_- := -x\mathbf{I}_{\{x<0\}}$. The quantity $-\min\{X,0\}$ is also called *shortfall* and its expectation is reminiscent of the *limited expected value function* known in the insurance literature. The Gain-Loss ratio is an acceptable index of performance, in the sense of [6].

- The mean-variance optimization problem can be replaced by the constrained maximization of m subject to the acceptability of X. This means that the inequality constrain $\rho(X) \leqslant \lambda$ is imposed on a different risk measure other than the standard deviation v, with $\lambda \geqslant 0$. In particular, ρ can be a coherent risk measure and one can equivalently refers to the maximization of the Risk-Adjusted Return on Capital (RAROC) $\frac{m}{\rho(X)}$, see for example [5], where X belongs to L^1 and is provided by admissible trading strategies. As a widespread coherent risk measure we refer to the expected shortfall or Average Value at Risk (AVaR), for which the statistical functional is $-\frac{1}{\alpha}\int_0^\alpha q_s^-(F)\mathrm{d}s$ with $X \sim F$. The integrand is the lower quantile $\inf\{x \in \mathbb{R} : F(x) \geqslant \alpha\}$ at the level $\alpha \in (0,1)$.

- Since practitioners use VaR as a risk measure, the performance ratio $\frac{m}{-q_\alpha^-(F)}$ at the level $\alpha \in (0,1)$ is of great importance and we include it in our analysis. Here $\mathrm{VaR}_\alpha(X) = -q_\alpha^-(F)$.

REMARK 1. Actually, the standard deviation of the random return is not a monetary risk measure at all in the sense of [13, Definition 4.1].

The statistical functionals corresponding to all the above performance measures are of the form $\theta(F) = \frac{\theta_1(F)}{\theta_2(F)}$, and are law-invariant since $\varphi(X) = \theta(F)$ for $X \sim F$. Thus their numerical value remain unchanged if computed for two different random returns X and Y having the same df. While the numerator of θ is always the mean, the denominator is a **risk functional**, i.e. a real-valued map from a subset $\mathrm{dom}(\theta_2) \subset \mathscr{D}$.

In this paper we consider only plug-in estimators $\theta(\mathbb{F}_n) = \frac{\theta_1(\mathbb{F}_n)}{\theta_2(\mathbb{F}_n)}$ based on n data points from historical returns. The AVaRR estimator is related to an L-statistic, where $X_{(i)}$ denotes the ith

Table 1: Estimators of Performance Ratios

	$\theta_1(\mathbb{F}_n)$	$\theta_2(\mathbb{F}_n)$
Sharpe ratio (SR)	$\frac{1}{n}\sum_{i=1}^n X_i$	$\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2\right)^{1/2}$
Gain-Loss ratio (GR)	$\frac{1}{n}\sum_{i=1}^n X_i$	$\frac{1}{n}\sum_{i=1}^n - \min\{X_i, 0\}$
AVaR ratio (AVaRR)	$\frac{1}{n}\sum_{i=1}^n X_i$	$-\frac{1}{n\alpha}\left(\sum_{i=1}^{\lfloor n\alpha\rfloor} X_{(i)} + X_{(\lfloor n\alpha\rfloor+1)}(n\alpha - \lfloor n\alpha\rfloor)\right)$
VaR ratio (VaRR)	$\frac{1}{n}\sum_{i=1}^n X_i$	$-X_{(\lfloor n\alpha\rfloor+1)}$

ordered statistics and $\lfloor y\rfloor$ denotes the integer part of $y \in \mathbb{R}$. The VaRR estimator is the sample quantile related to the empirical quantile function $\mathbb{F}_n^{-1}(\omega,\alpha) = X_{(i)}(\omega)\mathbf{I}_{\{\frac{i-1}{n}<\alpha\leqslant\frac{i}{n}\}}$ whose estimated counterpart is just $q_\alpha^-(\widehat{\mathbb{F}}_n)$.

From the perspective of asset pricing theory Cochrane and Saa-Request (2001) use the Sharpe ratio to spotlight the presence of portfolios with very high performance regarded as quasi-arbitrages or good-deals (see also [**?**] and [**?**]). To model efficiency in some instance or financial market stability (similar to the general equilibrium in economics) a basic assumption should be to rule out good-deals which in turn is equivalent to restrict Sharpe ratio. This way pricing kernels bounds are restricted if compared to those arising in a no-arbitrage framework. Unfortunately, a positive P&L such that $X \in L^1$ but whose variance is infinite produces zero Sharpe ratio still remaining an attractive investment alternative with positive underlying cash flow and large risk, actually an arbitrage.

4 Robustness from the Qualitative Viewpoint

In this section we will proof the following result:

Corollary 1. *Let X be a random return with distribution function F_X and assume $X \sim F \in \mathscr{D}^1 \cap$ dom(θ_2). Then the following performance measures have not a robust estimator $\widehat{\theta}_n$:*

- *The SR, for dom$(\theta_2) = \mathscr{D}^2$.*
- *The GR, for dom$(\theta_2) = \left\{F \in \mathscr{D} : \int x_- \mathrm{d}F < \infty\right\}$.*
- *The AVaRR, for dom$(\theta_2) = \left\{F \in \mathscr{D} : g \circ F \text{ induces } \mathrm{d}g(x) = \frac{1}{\alpha}\mathbf{I}_{\{0\leqslant x\leqslant\alpha\}}\mathrm{d}x\right\}$, where g is a continuous concave distortion.*
- *The VaRR, for dom$(\theta_2) = \{F \in \mathscr{D} : q_\alpha^+(F) = q_\alpha^-(F)\}$.*

Our plan is to first quickly review the definition of robustness, then to state a lemma which we use in the proof of Corollary 1.

4.1 Further Terminology

Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d. observations X_n with common df $F \in \mathscr{D}$. In ranking investment funds through performance ratios, one naturally requires that a contamination of the estimation procedure results in a small change of the resulting test statistics. If $\mathbf{y}$ is derived from the dataset $\mathbf{x}$ by largely modifying a small proportion of observations or by slightly modifying all of them, then the difference between the corresponding estimated empirical df's $\widehat{\mathbb{F}}_n$ should be small as well. The intuitive notion of robustness now calls for the continuity of θ provided that $\mathscr{D}$ is endowed with some metric measuring the distance between df's. When this is the case then for every sample size $n \in \mathbb{N}$ the value of $\theta(\widehat{\mathbb{F}}_n)$ will change not much. The Hampel's definition of robustness is valid for every dataset $(x_n)_{n\in\mathbb{N}}$, where x_n is a realization of X_n, and generalizes to any pair F, G of df's: the former corresponds to the true model of return X and the latter G corresponds to the contaminated model which deviates from the original in such a way a bias occurs in the estimation procedure. To formally represent the estimation procedure consider the composed maps

$$\mathbb{R}^{\mathbb{N}} \xrightarrow{\pi_{1,\ldots,n}} \mathbb{R}^n \xrightarrow{L} \mathscr{D} \xrightarrow{\theta} \mathbb{R},$$

where $\pi_{1,\ldots,n}(x_1, x_2, \ldots) = \mathbf{x} = (x_1, \ldots, x_n)$ is the projection of the first n data points and $L(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}_{\{x_i \leqslant x\}}$ is a linear combination of their point mass distributions. Thus, the estimator of θ is $\widehat{\theta}_n = h \circ L \circ \pi_{1,\ldots,n}$ and for every $n \in \mathbb{N}$ its value (estimate) at the corresponding dataset $\mathbf{x}$ is a measurable function $\widehat{\theta}_n(x_1, \ldots, x_n)$. The probability distribution of the estimator derived from F is $\mathsf{P}_F := \mathsf{P} \circ \widehat{\theta}_n^{-1}$, and a similar notation holds for the contaminated distribution G. Assuming θ is defined on some open subset of $\mathscr{D}$, the sequence of estimators $(\widehat{\theta}_n)_{n\in\mathbb{N}}$ is said to be **qualitatively robust** at F if for every $\varepsilon > 0$ there exist some $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for all $G \in \mathscr{D}$, which are in a neighborhood of F, we have $n \geqslant n_0$ and $d_{\text{Lévy}}(F, G) < \delta$ implies $d_{\text{Proh}}(\mathsf{P}_F, \mathsf{P}_G) < \varepsilon$. For the definition of the Lévy and Prohorov metrics see [14, Section 2.3]. When dealing with robust estimators of performance ratios we consequently have consistency, i.e. $\widehat{\theta}_n \xrightarrow{\mathsf{P}} \theta(F)$, independently of the i.i.d. sample **x** from F. In fact uniformly with probability 1 we have $\mathbb{F}_n \to F$, which in turn implies $d_{\text{Lévy}}(\mathbb{F}_n, F) \to 0$ with probability 1. If θ is continuous we can interchange $\theta(F)$ with the limit in probability $\mathbb{F}_n$ as $n \to \infty$, then the consistency is established. On the other hand, if consistency holds true then qualitative robustness is equivalent to continuity as stated in the classical Hampel's theorem, see [14, Proposition 2.20,Theorem 2.21].

4.2 Lack of Robustness

We study the qualitative robustness of each statistical functional $\theta(F) = \gamma(\theta_1(F), \theta_2(F))$ corresponding to the law-invariant performance ratios introduced so far, where $\gamma(x, y) = \frac{x}{y}$ is such that $y \neq 0$. Let $t : \mathscr{D} \to \mathbb{R}^2$ be defined by $t(F) = (\theta_1(F), \theta_2(F))$, so that $\theta = \gamma \circ t$. The aforementioned estimation procedure is equivalent to have plug-in estimators $\widehat{\theta}_n = \theta(\mathbb{F}_n)$ and $\widehat{\theta}_{n,i} = \theta_i(\mathbb{F}_n)$ each $i = 1, 2$. The following is an easy but useful result.

Lemma 1. *Let $\widehat{\theta}_{n,1}$ and $\widehat{\theta}_{n,2}$ be consistent estimators at the same F of the reward statistic $\theta_1(F)$ and the risk statistic $\theta_2(F)$, related to the law-invariant performance ratio $\varphi(X) = \theta(F)$ with $X \sim F$. If $\widehat{\theta}_{n,1}$ and $\widehat{\theta}_{n,2}$ are qualitatively robust at F, then $\widehat{\theta}_n$ is qualitatively robust at F.*

Proof. Assume $\widehat{\theta}_{n,1}$ and $\widehat{\theta}_{n,2}$ are robust estimators at F. Since they are consistent at F we have that $\widehat{\theta}_{n,1} \xrightarrow{\mathsf{P}} \theta_1(F)$ and $\widehat{\theta}_{n,2} \xrightarrow{\mathsf{P}} \theta_2(F)$ both imply $(\widehat{\theta}_{n,1}, \widehat{\theta}_{n,2}) \xrightarrow{\mathsf{P}} (\theta_1(F), \theta_2(F))$ by [25, Theorem 2.7]. The estimator $\widehat{\theta}_n$ of $\theta(F)$ is consistent at F by the continuity of γ [25, Theorem 2.3(ii)]. From the consistency of θ_1 and θ_2 together with the robustness hypotheses and [14, Theorem 2.21], it follows that they are continuous statistical functionals at F. The continuity of θ_1 and θ_2 is equivalent to that of the map t, thus θ is continuous too and the derived statistical functional is continuous at F. Another application of [14, Theorem 2.21] gives $\widehat{\theta}_n$ as a robust estimator of $\theta(F)$ and we are done. ∎

Before proving Corollary 1, let us provide some additional results. The risk functional corresponding to the AVaR can be represented as $\theta_2(F) = -\int_0^1 q_s^-(F)\mathrm{d}g(s)$, where $\mathrm{d}g(\cdot)$ is a probability measure on $[0, 1]$ related to the concave distortion g assumed to be continuous so that $g \circ F$ is a df. Thus, the AVaR is obtained by choosing $g(x) = \min\{\frac{x}{\alpha}, 1\}$, see Definition 4.48 and Lemma 4.69 in [13] for further details. When a decreasing density $f : [0, 1] \to [0, \infty)$ exists, the representation of AVaR is also given by $\mathrm{d}g(x) = f(x)\mathrm{d}x$; VaR can be given by the above representation through $g(x) = \mathbf{I}_{\{\alpha \leqslant x \leqslant 1\}}$; GR can be included in the class of RAROC measures with distortion $g(x) = x_-$, then the corresponding coherent risk measure is the expectation of the shortfall$-\min\{X, 0\}$.

Proof of Corollary 1. The sample mean $\widehat{\theta}_{n,1}$ and the sample standard deviation $\widehat{\theta}_{n,2}$ are not qualitatively robust (they are not continuous statistical functionals). Then by Lemma 1 their ratio is not a qualitatively robust estimator of the Sharpe ratio statistic. Replacing the standard deviation with m_- gives raise to a nonrobust estimator $\widehat{\theta}_{n,2}$ of the corresponding statistical functional $\theta_2(F)$, and again by Lemma 1 the GR statistic has not a qualitative robust estimator. Also, when $\theta_2(F)$ is $-q_\alpha^-(F)$ by Lemma 1 we do not gain any robustness of the performance ratio estimator, thought this risk functional has a qualitatively robust estimator $\widehat{\theta}_{n,2}$, see [14, Theorem 3.7] and [7, Theorems 3.4 and 3.5]. The risk functional corresponding to the AVaR has a spectral representation via the decreasing density $f(x) = \frac{1}{\alpha}\mathbf{I}_{\{0\leqslant x\leqslant\alpha\}}(x)$. This implies that $f \in L^q(0,1)$, $\frac{1}{p} + \frac{1}{q} = 1$ and $f(x) \neq 0$ in the neighborhood of 0. Thus, by [7, Propositions 2.2 and 3.6] and [7, Corollary 3.7] the estimator $\widehat{\theta}_{n,2}$ of such a risk functional is not qualitatively robust. ∎

REMARK 2. It is worth noting that all the above statistical functionals are defined at those F such that consistency of estimators is ensured. For example, the domain of the risk functional in the case of AVaRR and VaRR pertain to those df's for which the population quantile is uniquely determined. This provides the right link between the consistency of the estimator $\widehat{\theta}_{n,2}$ and the weak convergence $\widehat{F}_n \xrightarrow{\mathsf{P}} F$.

5 Conclusions

In judging fund management the performance attribution is not resistant to data contamination, then the estimated values of performance ratios do not fully reflect managerial skills. We show that a serious drawback is absence of qualitative robustness for Sharpe ratio, Gain-Loss ratio, Average-Value-at-Risk ratio and Value-at-Risk ratio. This suggests how not all the ratio statistics can be safely used in performance evaluation especially when alternative investment funds have to be ranked and the best has to be chosen.

References

[1] Ben-Tal, A., Nemirovski, A.: Robust Optimization - Methodology and Applications. Mathematical Programming Ser. B. **92**, pp. 453-480 (2002)

[2] Bernardo, A., Ledoit, O.: Gain, Loss and Asset Pricing. Journal of Political Economy. **108**, pp. 144-172 (2000)

[3] Broda, S.A., Paolella, M.S.: Expected Shortfall for Distributions in Finance. In: Statistical Tools for Finance and Insurance, pp. 57-99, Berlin (2011)

[4] Cavadini, F., Sbuelz, A., Trojani, F.: A Simplified Way of Incorporating Model Risk, Estimation Risk and Robustness in Mean Variance Portfolio Management. Working paper, Tilburg University, Tilburg, The Netherlands (2001)

[5] Cherny, A.: Weighted V@R and its Properties. Finance and Stochastics. **10**, pp. 367-393 (2006)

[6] Cherny, A., Madan, D.: New Measures for Performance Evaluation. Review of Financial Studies. **22**, pp. 2571-2606 (2009)

[7] Cont, R., Deguest, R., Scandolo, G.: Robustness and Sensitivity Analysis of Risk Measurement Procedures. Quantitative Finance. **10**(6), pp. 593-606 (2010)

[8] Costa, O.L.V., Paiva, A.C.: Robust Portfolio Selection using Linear-Matrix Inequality. Journal of Economic Dynamics & Control. **26**, pp. 889-909 (2002)

[9] Cowell, F.A., Victoria-Feser, M.-P.: Welfare Rankings in the Presence of Contaminated Data. Econometrica. **70**(3), pp. 1221-1233 (2002)

[10] DeMiguel, V., Nogales, F.J.: Portfolio Selection with Robust Estimation. Operations Research. **57**(3), pp. 560-577 (2009)

[11] Fama, E.F.: Efficient Capital Markets: a Review of Theory and Empirical Work. Journal of Finance. **25**, pp. 383-417 (1970)

[12] Fama, E.F.: Components of Investment Performance. Journal of Finance. **27**, pp. 551-567 (1972)

[13] Föllmer, H., Schied, A.: Stochastic Finance. De Gruyter, 3rd edition (2011)

[14] Huber, P.J., Ronchetti, E.M.: Robust Statistics. Wiley, 2nd edition (2009)

[15] Goldfarb, D., Iyengar, G.: Robust Portfolio Selection Problems. Mathematics of Operations Research. **28**, pp. 1-38 (2003)

[16] Jagannathan, R., Ma, T.: Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. Journal of Finance. **58**(4), pp. 1651-1684 (2003)

[17] Jorion, P.: Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection. Journal of Banking and Finance. **15**(3), pp. 717-727 (1991)

[18] Krätschmer, V., Schied, A., Zähle, H.: Qualitative and Infinitesimal Robustness of Tail-Dependent Statistical Functionals. Journal of Multivariate Analysis. **103**, pp. 35-47 (2012)

[19] Lauprete, G.J., Samarov, A.M., Welsch, R.E.: Robust Portfolio Optimization. Metrika **25**, pp. 139149 (2002)

[20] Lehmann, E.L.: Elements of Large-Sample Theory. Springer (2010)

[21] Lo, A.W.: The Statistics of Sharpe Ratios. Financial Analysts Journal. **58**(4), pp. 36-52 July/August (2002)

[22] Merton, R.C.: On Estimating the Expected Return on the Market: An Exploratory Investigation. Journal of Financial Economics. **8**, pp. 323-361 (1980)

[23] Perret-Gentil, C., Victoria-Feser, M.-P.: Robust Mean-Variance Portfolio Selection. FAME Research Paper 140. International Center for Financial Asset Management and Engineering, Geneva (2004)

[24] Serfling, R.J.: Approximation Theorems of Mathematical Statistics. Wiley, New York, NY (1980)

[25] van der Vaart, A.W.: Asymptotic Statistics. Cambridge Univ. Press, 8th printing (2007)

[26] van Zwet, W.: A Strong Law for Linear Functions of Ordered Statistics. Ann. Prob. **8**, pp. 986990 (1980)

[27] Welsch, R.E., Zhou, X.: Application of Robust Statistics to Asset Allocation Models. Revstat. Statist. J. **5**(1), pp. 97114 (2007)