Finance Research Letters 24 (2018) 105–112

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

A simulation comparison of risk measures for portfolio optimization

Marcelo Brutti Righi*, Denis Borenstein

School of Business, Federal University of Rio Grande do Sul, Washington Luiz 855, Porto Alegre, 90010-460, RS, Brazil

ARTICLE INFO

Article history:
Received 17 April 2017
Accepted 31 July 2017
Available online 2 August 2017

JEL classification:
C60
G10
G11

Keywords:
Risk measures
Loss-deviation
Portfolio selection
Simulation

ABSTRACT

In this paper, we compare risk measures regarding performance of optimal portfolio strategies. We consider eleven risk measures from different classes. In particular, we propose a formulation that generates from any loss measure, a deviation based on the dispersion of results worse than it, which leads to very interesting risk measures. We consider 198,000 portfolios composed by stocks of the U.S. equity market, considering different scenarios in a simulation framework. Results indicate there is no clearly dominant risk measure. Despite this lack of dominance, including deviation terms consistently exhibits advantages regarding performance.

1. Introduction

Portfolio optimization has been growing in importance among practitioners and researchers in finance since the classic mean-variance Markowitz's model. The research on modern portfolio theory is being concentrated on two main streams (Anagnostopoulos and Mamanis, 2010): (i) the incorporation of alternative risk measures, and (ii) the incorporation of real features in mathematical formulations. The present research work focuses on the former issue, experimentally comparing several risk measures when used in the portfolio selection.

The concept of risk in finance has been present since the beginning of financial transactions. The work of Markowitz (1952) pioneered the use of risk for decision-making. His modern portfolio theory is based on the dispersion or deviation of financial returns, measured by the variance or standard deviation as risk measure. In the past few decades, critical events and financial crises have turned attention to risk measures based on losses, such as Value at Risk (VaR) and Expected Shortfall (ES). Propelled by theoretical foundations of which properties a proper risk measure must fulfill, risk measures approaches can be broadly classified into two kinds: monetary or losses measures, which a stream of literature has begun with the seminal paper of Artzner et al. (1999), and deviation measures, with similar axiomatic theory started in Rockafellar et al. (2006). However, both types have drawbacks and limited applications (Sereda et al., 2010). Rachev et al. (2008) have proposed practical desirable properties that risk measures should exhibit for portfolio optimization.

Under this perspective, we expose a formulation that generates, from any loss measure, a deviation based on the dispersion of results worse than it. Thus, combining both leads to a loss-deviation risk measure. This kind of risk measure serves

* Corresponding author.
E-mail addresses: marcelo.righi@ufrgs.br, marcelobrutti@hotmail.com (M.B. Righi), denisb@ea.ufrgs.br (D. Borenstein).

as a more solid protection, once it yields higher values due to the penalty resulting from dispersion. We also present specific examples of risk measures that lie on the proposed approach. Some authors have proposed specific risk measures of this kind, beyond the intuitive mean plus standard deviation form the seminal Markowitz's approach. Ogryczak and Ruszczyński (1999) analyzed properties from the mean plus semi-deviation. Fischer (2003) and Chen and Wang (2008) considered combining the mean and semi-deviations at different powers to form a coherent risk measure. Furman and Landsman (2006) proposed a measure that weighs the mean and standard deviation in the truncated tail by VaR. Krokhmal (2007) extended the ES concept, for cases with higher moments with a relationship including deviation measures. Righi and Ceretta (2016) considered penalizing the ES by dispersion of results that represent losses exceeding the ES. Furman et al. (2017) penalize ES by the dispersion of tail based Gini measures. Righi (2017) exposes theoretical results of risk measures composed by losses and deviations, but the author restricts to specific classes and do not propose practical examples and illustrations. Thus, we contribute to existing knowledge because we expose a new approach that allows to directly construct from any chosen loss measure. Rockafellar et al. (2006) proved that it is possible to construct a lower range dominated generalized deviation measure from a loss measure. Nonetheless, our construction is based on penalization, rather than a conversion, as is their case.

Ortobelli et al. (2005) stated several risk measures are equivalent to generate the same solution when used in optimization problems for a given investor's category. However, a comparison between several allocation problems using theoretically equivalent risk measures led to significant differences in the portfolio choices (Giacometti and Ortobelli, 2004; Brandtner et al., 2017). Although these inconsistencies can depend on other issues, such as heavy tails or asymmetries in the data, different risk measures can penalize/favor properties that can deteriorate/improve the performance of the optimization model. Based on this, it is highly relevant to compare the performance of different measures in different classes of risk measures towards answering the following questions: (i) Which risk measure is more appropriate for what type of investor's strategy; (ii) What are the advantages/disadvantages of using a risk measure over another; and (iii) What portfolio size and investment horizon affect performance.

Several risk measures of varying complexity and possibly solutions may be suitable for use in a particular portfolio selection problem. The choice of the risk measure can be a rather complicated task for non-expert users, leading to the choice of the easiest ones to compute or the one that leads to linearizable optimization formulations (easier to solve in large-scale portfolio selections problems). A systematized comparison of different risk measures was recognized and has been carried from a theoretical point of view using their mathematical properties (see, for instance, Ogryczak and Ruszczyński (1999); Brandtner (2013); Emmer et al. (2015); Pichler (2017)). Some researchers presented empirical comparisons of risk measures regarding portfolio optimization, such as Alexander and Baptista (2002); 2004); Ortobelli et al. (2005); DeMiguel et al. (2009); Pflug et al. (2012), comparing few specific risk measures with the traditional mean-variance approach in small case examples. The main finding of these studies is the lack of consensus about a definitive or superior risk measure. Based on this finding, there is a clear gap in the literature for a systematic empirical comparison, simultaneously using a large number of risk measures and comparing the performance of them using several scenarios, with different number of assets, optimization strategies, and data sample sizes. The relevance of extensive empirical analyses was demonstrated in the systemic risk measure area by Kleinow et al. (2017).

In this paper, we present a simulation study comparing eleven risk measures, since the standard deviation and VaR, to some coherent, convex, and loss-deviation measures. We consider both risk minimization and maximization of the ratio between expected return and risk in a Sharpe ratio fashion. For the empirical analysis, we use daily data from the U.S. equity market, composing portfolios with different numbers of assets in a procedure inspired by Monte Carlo simulation. We verify the performance of such portfolios in distinct out-sample investment horizons using, as metrics, beyond descriptive statistics of returns series, all risk measures present in the study, as well as the ratio between average return to them. Results are analyzed using statistical tests of difference and advantage of ranks between obtained portfolios.

Our study contributes to the existing literature, offering a comprehensive empirical analysis of different risk measures. To the best of our knowledge, there is no extensive comparison of risk measures for portfolio optimization, considering loss-deviation measures. Our work emphasizes the practical issues in risk measuring for the portfolio selection problem, proposing a simulation study with large sample sizes and a wide range of problems. As a consequence, we can satisfactorily answer the abovementioned questions.

The remaining of this paper is structured as follows: Section 2 presents the eleven risk measures used in the comparison. Section 3 describes the employed methodological procedures in our experiments. The results are presented and analyzed in Section 4. Finally, Section 5 concludes the paper.

2. Selected risk measures

Consider the random result X of any asset ($X \geq 0$ is a gain, $X < 0$ is a loss) defined in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $E[X]$ is the expected value of X under $\mathbb{P}$. All equalities and inequalities are considered almost surely in $\mathbb{P}$. F_X is the probability function of X and its inverse is F_X^{-1}, defined as $F_X^{-1}(\alpha) = \inf\{x : F_X(x) \geq \alpha\}$. Let $L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$ be the space of random

The measures proposed in Righi (2017) have the generic form $\rho + \mathcal{D}$, where ρ is a coherent risk measure in the sense of Artzner et al. (1999), and $\mathcal{D}$ is a generalized deviation measure, as proposed by Rockafellar et al. (2006). A link is made through an axiom of Limitedness, stating this set of measures is a sub-class of coherent risk measures. In this paper, our focus is more on practical meaning than theoretical properties. In this sense, we consider dispersion measured by the p-norm semi-deviation of results that represent losses greater than ρ as the conform $\mathcal{D}(X) = \|(X - \rho^*(X))^-\|_p, \rho^*(X) = -\rho(X)$. Since the objective is to penalize the risk measured by ρ, we only consider the dispersion of results that represent losses greater than this value. The role of the minus sign is simply an adjustment to place ρ at the same level of X, because the former represents losses and the latter the results of an asset. In this way, given any loss risk measure ρ we introduce one new conform $\rho(X) + \beta\|(X - \rho^*(X))^-\|_p, 0 \leq \beta \leq 1$, which can be understood as a loss penalized by the dispersion of results worse than this conform. The role of β is to choose the proportion of dispersion that has to be included. Thus, it functions similar to an aversion term. Hence, such a risk measure can be thought of as being more comprehensive because it yields more solid protection due to the penalty by dispersion.

We evaluate, in this paper, traditional risk measures from the literature and their loss-deviation companions to have a more complete analysis. Thus, we consider eleven risk measures for comparison in our research, for $0 \leq \alpha,\ \beta \leq 1$, as follows:

1. Standard deviation: $StD(X) = \|(X - E[X])\|_p$;
2. Value at risk: $VaR^\alpha = -F_X^{-1}(\alpha)$;
3. Expected loss: $EL(X) = E[-X]$;
4. Expected loss deviation: $ELD(X) = E[-X] + \beta\|(X - E[X])^-\|_p$;
5. Expected shortfall: $ES^\alpha(X) = E[-X|X \leq F_X^{-1}(\alpha)]$;
6. Shortfall deviation risk: $SDR^\alpha(X) = ES^\alpha(X) + \beta\|(X - ES^{*,\alpha}(X))^-\|_p$;
7. Expectile value at risk: $EVaR^\alpha = -\arg\min_\gamma E[(\alpha - \mathbf{1}_{X\leq\gamma})(X-\gamma)^2]$;
8. Deviation expectile value at risk: $DEVaR^\alpha(X) = EVaR^\alpha(X) + \beta\|(X - EVaR^{*,\alpha}(X))^-\|_p$;
9. Entropic: $ENT^\theta(X) = \frac{1}{\theta}\log\left(E[e^{-\theta X}]\right), \theta \geq 0$;
10. Deviation entropic: $DENT^\theta(X) = ENT^\theta(X) + \beta\|(X - ENT^{*,\theta}(X))^-\|_p$;
11. Maximum loss: $ML(X) = -\inf X = \sup -X$.

The rationale for selection has been to consider risk measures among the most popular and widely used and their loss-deviation counterpart. These measures reflect different classes and present distinct complexity. Our first choice is the usual p-norm standard deviation (StD), to represent the Markowitz classical approach. Our second choice is the canonical tail risk measure, the well-known and largely used VaR. Focusing on coherent risk measures, we also consider the Expected Loss (EL), which generates the loss-deviation risk measure Expected Loss Deviation (ELD). This risk measure was studied by Ogryczak and Ruszczyński (1999) and Fischer (2003). We also consider the ES, proposed by Acerbi and Tasche (2002), which represents the expected value of a loss, given it is beyond the α-quantile of interest, directly linked to the VaR concept. It generates the loss-deviation risk measure proposed in Righi and Ceretta (2016), the Shortfall Deviation Risk (SDR). Our next choice regards a risk measure that has recently gained more attention, the Expectile Value at Risk (EVaR). This measure is linked to the concept of expectile. Bellini et al. (2014) and Ziegel (2016) study EVaR, verifying it is the only coherent risk measure beyond the mean loss that possesses the property of elicitability, which allows a function to have its forecasts evaluated. Righi (2017) consider the loss-deviation risk measure Deviation EVaR (DEVaR). We also consider a convex, but not coherent risk measure, studied by Föllmer and Schied (2002), the Entropic (ENT). This risk measure leads to the loss-deviation Deviation ENT (DENT). Finally, we also consider the Maximum Loss (ML), which is the more conservative risk measure, since it represents the most extreme loss of a financial position.

3. Simulation experiment

To verify the performance of distinct risk measures for portfolio optimization, we consider daily log-returns of 2285 stocks from the U.S. equity market, in the period between January 2010 and December 2015, totalizing 1547 observations. This time interval considers a period after the sub-prime financial crisis, contemplating both calm and turbulent market conditions. It is an interesting environment for financial diversification. We consider stocks negotiated in the whole period. Our experiment simulates an investor with 2285 assets available, but desires a limited number of them in the composition of his/her portfolio. Descriptive results, available upon request, indicate the return series exhibits known stylized facts of financial data, such as zero mean, distinct dispersion, negative asymmetry, and heavy tails. This configures, among other characteristics, an uncertain typical scenario of equity markets, highlighting the need for proper risk management, such as risk diversification in different portfolio strategies.

We consider two distinct strategies, risk minimization and maximization of the ratio between expected returns (mean) mization models can be formulated as follows:

Minimization:

$$\min_{W\in\mathbb{R}^k} \rho\left(\sum_{i=1}^{k} w_i X_i\right)$$

st

$$\sum_{i=1}^{k} w_i = 1$$

$$w_i \geq 0,\ i = 1, \ldots, k$$

Maximization:

$$\max_{W\in\mathbb{R}^k} \frac{E\left[\sum_{i=1}^{k} w_i X_i\right]}{\rho\left(\sum_{i=1}^{k} w_i X_i\right)}$$

st

$$\sum_{i=1}^{k} w_i = 1$$

$$w_i \geq 0,\ i = 1, \ldots, k \quad (1)$$

Both problems do not allow for short selling and requires all capital will be allocated, represented by constraint (1). These are very common constraints in this kind of problem.

Risk measures are estimated by the empirical method, known as historical simulation (HS), which is a nonparametric method that creates no assumptions about the data and is the most extensively used method in academic studies and in the financial industry, conform exposed in Pérignon and Smith (2010). For this approach, formulations of Section 2 are computed considering the empirical distribution. Regarding parameters, we choose $p = 2,\ \alpha = 0.05,\ \beta = 1$. Here, we represent financial positions with defined second moment, a quantile of interest used on financial studies and risk management, and a risk aversion coefficient that incorporates all the deviation terms. This configuration is related to the concept of (semi-) standard deviation, which has a strong intuitive financial meaning. For ENT and DENT, we choose $\theta = 1$. For robustness matters, we also consider other values for these parameters, but there is no qualitatively change in the obtained results.

We consider portfolios composed by 4, 16, and 64 assets. For each asset number, we obtain the optimal weights on both strategies for each of the eleven risk measures, maintaining the strategy over 125, 250, and 500 trading days. We consider in parallel results with re-balance of portfolios, but there are no significant changes on the obtained results. We keep them without re-balance. We also consider distinct investment horizons. The remaining previous 1047 $(1547 - 500)$ days are used in the optimization step. This sample size, which represents approximately 4 years of daily observations, is indicated in studies that compare risk measure estimators, such as Kuester et al. (2006), because it produces lower estimation errors. We have 9 scenarios (3 asset numbers times 3 horizons). We also consider other scenarios with different number of assets and horizons, but the results have no qualitative change. We keep our original results. In each scenario, we replicate the procedure 1000 times with random samples from the 2285 stocks. Although similar to a Monte Carlo simulation, we need not specify any data generation process, which provides a more realistic empirical analysis.

For every portfolio in each scenario for both strategies, we compute the following performance metrics: maximum, mean, skewness and kurtosis, as well as the eleven risk measures and the ratio of obtained average return (mean) to each of the eleven risk measures. Note that standard deviation and the minimum are also contemplated by StD and ML, respectively. We have 26 criteria for performance. We then compute the average of such performance metrics over the 1000 replicates to rank the best portfolio (1 for the best, 11 for the worst). Obviously, larger values of mean, skewness, maximum, and ratios of mean to risk measures are desirable, while smaller values for kurtosis and risk measures are preferable. Then, for each scenario, we compute the average rank over the 26 criteria for each portfolio to have an aggregated value representing the overall performance. We have, for each risk measure, an average of ranks for every scenario. Original results for computed criteria and ranks are available upon request.

On such average ranks, we apply Kruskal-Walis tests to verify if there is jointly difference in their distributions. We also apply bi-tailed paired Wilcoxon tests to verify if differences in the average ranks for portfolios are obtained with a specific risk measure regarding the strategy (minimization of risk and maximization of return per risk). We apply bi-tailed and left one-tailed Wilcoxon paired tests to verify if there is, respectively, difference in means and advantage, in the sense of small ranks, in the performance of portfolios. We choose such statistical tests due to their non-parametric nature, which makes minimal assumptions about data structure. We consider a significance level of 10% in the analyses. Since such procedures are conducted over small samples, we repeated them on the ranks, considering all 26 performance criteria, instead of their averages, but we have not find enough differences from the original results. Due to parsimony, we keep the initial results, but the alternative ones are available upon request. These procedures are also conducted to verify differences regarding scenarios, i.e., for each scenario, we apply the same tests over all portfolios obtained from distinct risk measures. This is done to observe the robustness of results regarding our choice of scenarios.

4. Results

Tables 1 and 2 present the average ranks of portfolios obtained with all the risk measures for the minimization and maximization strategies, respectively, and scenarios. We note different risk measures emerge as those with the best ranks for

Table 1
Average of ranks for minimization strategy considering distinct risk measures.

	Risk measure										
Scenario	StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
4_500	6.54	5.07	6.79	6.18	7.79	4.86	7.36	6.79	5.32	5.21	4.11
4_250	5.07	4.50	4.07	7.29	7.75	6.46	5.64	6.86	5.75	6.96	5.64
4_125	6.61	5.18	7.04	5.25	5.86	7.04	6.71	6.57	5.39	5.25	5.11
16_500	6.64	6.64	6.29	5.32	6.25	6.71	4.75	5.93	6.39	4.46	6.61
16_250	6.61	6.96	6.93	5.00	5.61	5.71	4.50	5.32	7.50	5.64	6.21
16_125	5.61	6.29	6.57	5.18	6.43	6.36	5.93	5.96	6.93	4.75	6.00
64_500	5.96	6.32	6.61	5.82	5.79	5.79	5.71	5.64	6.79	5.50	6.07
64_250	6.14	6.43	6.07	5.36	6.46	5.89	6.04	6.04	5.79	5.46	6.32
64_125	5.96	6.14	6.50	5.79	6.07	5.75	6.00	5.93	6.36	5.64	5.86
Average	6.13	5.95	6.32	5.69	6.44	6.06	5.85	6.12	6.25	5.43	5.77

Kruskal-Walis p-value = 0.09

Table 2
Average ranks for maximization strategy considering distinct risk measures.

	Risk measure										
Scenario	StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
4_500	5.14	5.86	6.32	4.96	6.68	6.71	6.54	5.93	5.89	5.00	6.96
4_250	5.36	5.25	7.61	6.25	6.71	4.50	7.25	5.61	5.86	5.07	6.54
4_125	4.82	5.43	8.71	5.82	6.29	5.39	6.50	5.46	6.64	5.11	5.82
16_500	5.18	5.50	6.18	6.32	6.07	6.54	6.21	6.43	4.93	5.93	6.71
16_250	5.11	5.82	8.14	5.93	7.25	6.29	5.43	4.46	5.39	5.00	7.18
16_125	5.25	6.32	6.46	6.29	6.82	4.96	6.79	4.93	4.96	6.29	6.93
64_500	5.79	6.46	6.36	6.04	5.61	5.75	6.79	5.93	5.36	5.68	6.25
64_250	5.54	6.21	6.71	6.18	5.61	6.11	5.64	6.57	4.89	6.00	6.54
64_125	6.14	6.25	6.00	6.21	6.18	6.11	5.61	6.07	5.96	5.61	5.86
Average	5.37	5.90	6.94	6.00	6.36	5.82	6.31	5.71	5.54	5.52	6.53

Kruskal-Walis p-value = 0.00

Table 3
Paired Wilcoxon between strategies.

	Risk measure										
Measure	StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
p-value	**0.03**	0.91	0.41	0.50	0.72	0.82	0.16	0.15	0.13	1.00	**0.01**

Bold values are significant at 10% level

strategies, namely for StD (improve in maximization) and ML (better in minimization). This result indicates a portfolio based on determined risk measure has similar performance for both risk minimization and maximization of return adjusted by risk. Distinct investor strategies can be attended with the same risk measures.

We present in Tables 4, 5, 6, and 7 Wilcoxon paired tests for difference (bi-tailed) and advantage (left one-tailed) in ranks of portfolios composed by distinct risk measures. Results for risk minimization exhibit few significant rejections of the null hypothesis, indicating similarity on performance. This can be due to the fact such strategy is naïve, because it neglects profitability, focusing on protection and all alternatives for risk we consider have good theoretical properties. Some risk measures exhibit better performance for obtained portfolios, such as DENT, ELD, and ML. Regarding maximization of the ratio between return and risk, we observe more significant rejections, indicating a larger overall number of distinctions in performance. Since this is a more adequate strategy, because it balances the two most relevant variables in financial decision, results are more interesting. Here, risk measures with best performance are StD, ENT, DENT, SDR, ELD, and DEVaR, since they either have advantage or do not have disadvantage against a reasonable number of competitors.

It is possible to state there is not a dominant risk measure, ratifying the findings of previous theoretical and empirical comparisons. This can be explained because each risk measure has its own peculiarities and interpretation. One can conclude distinct financial situations require particular risk measures that are more suited. We note that VaR, a non-convex (coherent) risk measure, does not exhibit good performance for our optimization problems. ENT and its loss-deviation counterpart DENT, which are convex but not coherent and less used in practical matters, present good performance. StD has competitive performance, even being the simplest measure we consider. Despite this lack of dominance, loss-deviation risk measures

Table 4
Bi-tailed paired Wilcoxon tests for distinct risk measures ranks – minimization strategy.

	Risk measure										
Measure	StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
StD	1.00	0.83	0.36	0.13	0.65	0.82	0.95	0.73	0.65	0.10	0.73
VaR	0.83	1.00	0.65	0.43	0.65	0.81	0.73	1.00	0.19	0.25	0.13
EL	0.36	0.65	1.00	0.12	0.65	0.36	0.25	0.18	0.91	0.13	0.16
ELD	0.13	0.43	0.12	1.00	**0.01**	0.43	0.57	**0.05**	0.20	0.14	0.73
ES	0.65	0.65	0.65	0.01	1.00	0.44	**0.05**	**0.07**	0.82	**0.01**	0.30
SDR	0.82	0.81	0.36	0.43	0.44	1.00	0.36	0.63	0.57	0.16	0.43
EVaR	0.95	0.73	0.25	0.57	**0.05**	0.36	1.00	0.62	0.50	0.25	0.73
DEVaR	0.73	1.00	0.18	**0.05**	**0.07**	0.63	0.62	1.00	0.91	**0.04**	0.73
ENT	0.65	0.19	0.91	0.20	0.82	0.57	0.50	0.91	1.00	**0.04**	0.07
DENT	**0.10**	0.25	0.13	0.14	**0.01**	0.16	0.25	**0.04**	**0.04**	1.00	0.41
ML	0.73	0.13	0.16	0.73	0.30	0.43	0.73	0.73	**0.07**	0.41	1.00

We have that cell (i, j) represents advantage of risk measure in line i against the competitor on column j. Bold values are significant at 10% level.

Table 5
Left one-tailed paired Wilcoxon tests for distinct risk measures ranks – minimization strategy.

	Risk measure										
Measure	StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
StD	1.00	0.64	0.18	0.95	0.33	0.63	0.57	0.67	0.33	0.96	0.67
VaR	0.42	1.00	0.33	0.82	0.33	0.64	0.67	0.50	0.10	0.90	0.95
EL	0.85	0.71	1.00	0.95	0.71	0.85	0.90	0.93	0.46	0.95	0.94
ELD	**0.06**	0.21	**0.06**	1.00	**0.01**	0.21	0.29	**0.03**	0.10	0.95	0.37
ES	0.71	0.71	0.33	1.00	1.00	0.82	0.98	0.97	0.63	1.00	0.88
SDR	0.41	0.41	0.18	0.82	0.22	1.00	0.85	0.72	0.29	0.94	0.82
EVaR	0.48	0.37	0.12	0.75	0.03	0.18	1.00	0.31	0.25	0.90	0.36
DEVaR	0.37	0.54	0.09	0.98	0.04	0.32	0.74	1.00	0.46	0.99	0.67
ENT	0.71	0.92	0.59	0.92	0.41	0.75	0.79	0.59	1.00	0.99	0.97
DENT	**0.05**	0.12	**0.06**	**0.07**	**0.00**	**0.08**	0.12	**0.02**	**0.02**	1.00	0.20
ML	0.37	**0.06**	**0.08**	0.67	0.15	0.21	0.69	0.37	**0.04**	0.83	1.00

We have that cell (i, j) represents advantage of risk measure in line i against the competitor on column j. Bold values are significant at 10% level.

Table 6
Bi-tailed paired Wilcoxon tests for distinct risk measures ranks – maximization strategy.

	Risk measure										
Measure	StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
StD	1.00	**0.01**	**0.01**	**0.01**	**0.02**	0.25	**0.02**	0.16	0.65	0.59	**0.01**
VaR	**0.01**	1.00	**0.03**	0.81	0.25	0.73	0.20	0.72	0.30	**0.04**	**0.03**
EL	**0.01**	**0.03**	1.00	**0.03**	0.12	**0.05**	0.25	**0.03**	**0.00**	**0.00**	0.43
ELD	**0.01**	0.81	**0.03**	1.00	0.30	0.51	0.26	0.30	0.20	**0.02**	**0.06**
ES	**0.02**	0.25	0.12	0.30	1.00	0.30	0.95	0.20	**0.02**	**0.03**	0.41
SDR	0.25	0.73	**0.05**	0.51	0.30	1.00	0.43	1.00	0.53	0.30	**0.02**
EVaR	**0.02**	0.20	0.25	0.26	0.95	0.43	1.00	0.10	0.04	**0.03**	0.65
DEVaR	0.16	0.72	**0.03**	0.30	0.20	1.00	0.10	1.00	0.73	0.37	**0.02**
ENT	0.65	0.30	**0.00**	0.20	**0.02**	0.53	**0.04**	0.73	1.00	1.00	**0.03**
DENT	0.59	**0.04**	**0.00**	**0.02**	**0.03**	0.30	**0.03**	0.37	1.00	1.00	**0.00**
ML	**0.01**	**0.03**	0.43	**0.06**	0.41	**0.02**	0.65	**0.02**	**0.03**	**0.00**	1.00

We have that cell (i, j) represents advantage of risk measure in line i against the competitor on column j. Bold values are significant at 10% level.

that, in such an application, the goal is more on risk diversification than protection for larger but rare losses. The inclusion of deviation terms can be justified.

Concerning the scenarios we consider, we repeat all statistical tests considering the ranks of all risk measures, i.e., by transposing the content in Tables 1 and 2. However, we find no rejection of the null hypotheses. This result highlights that ranks concisely do not change from one scenario to another. This implies our obtained results are robust to the choice

Table 7
Left one-tailed paired Wilcoxon tests for distinct risk measures ranks – minimization strategy.

Measure	Risk measure StD	VaR	EL	ELD	ES	SDR	EVaR	DEVaR	ENT	DENT	ML
StD	1.00	**0.01**	**0.00**	**0.01**	**0.01**	0.12	**0.01**	**0.08**	0.33	0.30	**0.01**
VaR	0.99	1.00	**0.01**	0.41	0.12	0.67	0.10	0.68	0.88	0.99	**0.01**
EL	1.00	0.99	1.00	0.99	0.95	0.98	0.90	0.99	1.00	1.00	0.82
ELD	1.00	0.64	**0.02**	1.00	0.15	0.78	0.13	0.88	0.92	0.99	**0.03**
ES	0.99	0.90	**0.06**	0.88	1.00	0.88	0.48	0.92	0.99	0.99	0.20
SDR	0.90	0.37	**0.03**	0.26	0.15	1.00	0.21	0.54	0.78	0.88	**0.01**
EVaR	0.99	0.92	0.12	0.89	0.57	0.82	1.00	0.96	0.99	0.99	0.33
DEVaR	0.94	0.36	**0.01**	0.15	0.10	0.50	**0.05**	1.00	0.67	0.84	**0.01**
ENT	0.71	0.15	**0.00**	0.10	**0.01**	0.26	**0.02**	0.37	1.00	0.54	**0.01**
DENT	0.74	**0.02**	**0.00**	**0.01**	**0.01**	0.15	**0.01**	0.19	0.50	1.00	**0.00**
ML	1.00	0.99	0.21	0.98	0.83	0.99	0.71	0.99	0.99	1.00	1.00

We have that cell (*i*, *j*) represents advantage of risk measure in line *i* against the competitor on column *j*. Bold values are significant at 10% level.

5. Conclusion

In this paper, we compare risk measures regarding performance of optimal portfolio strategies. We consider both risk minimization and maximization of the ratio between mean and risk, with eleven risk measures. We propose a formulation that generates from any loss measure, a deviation based on the dispersion of results worse than it. Combining both we have the concept of loss-deviation risk measures. We generate 198,000 portfolios composed by stocks of the U.S. equity market, considering different number of assets and investment horizons within a simulation framework.

Results indicate the eleven compared risk measures produce different performances when employed in portfolio optimization. However, there is no clearly dominant risk measure, but some patterns arise, such as more differences in maximization than minimization strategies. Despite this lack of dominance, loss-deviation risk measures consistently exhibit advantage regarding performance, having good practical results in our context. This result advocates in favor of more complete risk measures, with the penalization by dispersion of outcomes worse than some loss bound.

Overall, the final choice of a risk measure cannot be only defined based on performance comparison, but also considering other relevant issues such as the complexity of the portfolio optimization model when using a specific risk measure. Multi-objective optimization (MOO) has become the state-of-the-art to represent portfolio selection problems (Zopounidis et al., 2015). These models are difficult to solve in portfolio selection, mainly when several real features are incorporated in their mathematical formulations. As some risk measures can be linearizable, this can be seen as a very good advantage, implying much quicker solution algorithms for MOO. Future research is proceeding to compare risk measures in MOO portfolio optimization.

References

Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. J. Bank. Finance 26, 1487–1503.
Alexander, G.J., Baptista, A.M., 2002. Economic implications of using a mean-var model for portfolio selection: a comparison with mean-variance analysis. J. Econ. Dyn. Control 26 (7), 1159–1193.
Alexander, G.J., Baptista, A.M., 2004. A comparison of VaR and CVar constraints on portfolio selection with the mean-variance model. Manage. Sci. 50 (9), 1261–1273.
Anagnostopoulos, K., Mamanis, G., 2010. A portfolio optimization model with three objectives and discrete variables. Comput. Oper. Res. 37 (7), 1285–1297.
Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Math. Finance 9 (3), 203–228.
Bellini, F., Klar, B., Müller, A., Rosazza Gianin, E., 2014. Generalized quantiles as risk measures. Insur. Math. Econ. 54 (2014), 41–48.
Brandtner, M., 2013. Conditional value-at-risk, spectral risk measures and (non-) diversification in portfolio selection problems–a comparison with mean–variance analysis. J. Bank. Finance 37 (12), 5526–5537.
Brandtner, M., Kürsten, W., Rischau, R., 2017. Entropic risk measures and their comparative statics in portfolio selection: coherence vs. convexity. Eur. J. Oper. Res. in press, –.
Chen, Z., Wang, Y., 2008. Two-sided coherent risk measures and their application in realistic portfolio optimization. J. Bank. Finance 32 (12), 2667–2673.
DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: how inefficient is the 1/n portfolio strategy? Rev. Financ. Stud. 22 (5), 1915–1953.
Emmer, S., Kratz, M., Tasche, D., 2015. What is the best risk measure in practice? a comparison of standard measures. J. Risk 18 (2), 31–60.
Fischer, T., 2003. Risk capital allocation by coherent risk measures based on one-sided moments. Insur. Math. Econ. 32, 135–146.
Föllmer, H., Schied, A., 2002. Convex measures of risk and trading constraints. Finance Stochastics 6, 429–447.
Furman, E., Landsman, Z., 2006. Tail variance premium with applications for elliptical portfolio of risks. ASTIN Bull. 36 (2), 433–462.
Furman, E., Wang, R., Zitikis, R., 2017. Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. J. Bank. Finance 83, 70–84.
Giacometti, R., Ortobelli, S., 2004. Risk Measures for Asset Allocation Models. In: Szegö, G. (Ed.), Risk Measure for the 21st Century. Wiley & Son, Chischester, pp. 69–87.
Kleinow, J., Moreira, F., Strobl, S., Vähämaa, S., 2017. Measuring systemic risk: a comparison of alternative market-based approaches. Finance Res. Lett 21, 40–46.

Ogryczak, W., Ruszczyński, A., 1999. From stochastic dominance to mean-risk models: semideviations as risk measures. Eur. J. Oper. Res. 116 (1), 33–50.
Ortobelli, S., Rachev, S.T., Stoyanov, S., Fabozzi, F.J., Biglova, A., 2005. The proper use of risk measures in portfolio theory. Int. J. Theor. Appl. Finance 8 (08), 1107–1133.
Pérignon, C., Smith, D.R., 2010. The level and quality of value-at-risk disclosure by commercial banks. J. Bank. Finance 34 (2), 362–377.
Pflug, G.C., Pichler, A., Wozabal, D., 2012. The 1/N investment strategy is optimal under high model ambiguity. J. Bank. Finance 36 (2), 410–417.
Pichler, A., 2017. A quantitative comparison of risk measures. Ann. Oper. Res 254 (1–2), 251–275.
Rachev, S., Ortobelli, S., Stoyanov, S., Fabozzi, F.J., Biglova, A., 2008. Desirable properties of an ideal risk measure in portfolio theory. Int. J. Theor. Appl. Finance 11 (01), 19–54.
Righi, M., 2017. A composition between risk and deviation measures. Work. Pap. https://arxiv.org/abs/1511.06943.
Righi, M., Ceretta, P., 2016. Shortfall deviation risk: an alternative to risk measurement. J. Risk 19 (2), 81–116.
Rockafellar, R., Uryasev, S., Zabarankin, M., 2006. Generalized deviations in risk analysis. Finance Stochastics 10, 51–74.
Sereda, E.N., Bronshtein, E.M., Rachev, S.T., Fabozzi, F.J., Sun, W., Stoyanov, S.V., 2010. Distortion risk measures in portfolio optimization. In: Handbook of Portfolio Construction. Springer, pp. 649–673.
Ziegel, J., 2016. Coherence and elicitability. Math. Finance 26 (4), 901–918.
Zopounidis, C., Galariotis, E., Doumpos, M., Sarri, S., Andriosopoulos, K., 2015. Multiple criteria decision aiding for finance: an updated bibliographic survey. Eur. J. Oper. Res. 247 (2), 339–348.