Quantitative Finance, 2021
Vol. 21, No. 2, 243–261, https://doi.org/10.1080/14697688.2020.1820072

Tail risks in large portfolio selection: penalized quantile and expectile minimum deviation models

R. GIACOMETTI †‡, G. TORRI †‡ and S. PATERLINI *§

†Department of Management, Economics and Quantitative Methods, University of Bergamo, Via dei Caniana, 2, 24127 Bergamo, Italy
‡Department of Finance, VŠB-TU Ostrava, Sokolská třída 33, Ostrava 701 21, Czech Republic
§Department of Economics and Management, University of Trento, via Inama 5, I-38122 Trento, Italy

(*Received 2 March 2020; accepted 28 August 2020; published online 27 October 2020*)

Accurate estimation and optimal control of tail risk is important for building portfolios with desirable properties, especially when dealing with a large set of assets. In this work, we consider optimal asset allocation strategies based on the minimization of two asymmetric deviation measures, related to quantile and expectile regression, respectively. Their properties are discussed in relation with the 'risk quadrangle' framework introduced by Rockafellar and Uryasev [The fundamental risk quadrangle in risk management, optimization and statistical estimation. *Surv. Oper. Res. Manag. Sci.*, 2013, **18**(1–2), 33–53], and compared to traditional strategies, such as the mean–variance portfolio. In order to control estimation error and improve the out-of-sample performance of the proposed models, we include ridge and elastic-net regularization penalties. Finally, we propose quadratic programming formulations for the optimization problems. Simulations and real-world analyses on multiple datasets allow to discuss pros and cons of the different methods. The results show that the ridge and elastic-net allocations are effective in improving the out-of-sample performance, especially in large portfolios, compared to the un-penalized ones.

Keywords: Tail risk; Expectiles; Quantiles; Regularization; Portfolio optimization

1. Introduction

Starting from Markowitz (1952), the identification of optimal portfolio allocations in terms of risk and returns has been a primary focus for both practitioners and academics. In such cornerstone papers, the author considers an investor that seeks high returns, but is averse to *risk* in the form of variance. This model, which is consistent with an investor characterized by a quadratic utility function, has a major drawback: it symmetrically penalizes deviations from the expected returns. For this reason, several authors proposed alternative measures to assess risk such as semi-deviation, value at risk (*V aR*), expectile (Bellini and Di Bernardino 2017) and conditional value at risk (*CVaR*) to take into account tail risk. In 1999, Artzner *et al.* (1999) introduced the axiomatic definition of *coherent risk measures*. According to this definition, not all the risk measures used in the literature can be considered coherent (e.g. variance, standard deviation and *VaR* are not). Moreover, some of them are classified as deviation measures, rather than risk measures (e.g. variance and standard deviation) (Rockafellar *et al.* 2008).

The relationship between risk and deviation measures has been further investigated by Rockafellar and Uryasev (2013), who proposed the 'risk quadrangle', a framework that links risk and deviation measures, discussing their estimation procedure based on the minimization of an error measure. Such a framework nests several well-known measures in the literature, and it can be related to common portfolio allocation strategies such as the mean–variance and the mean–*CV aR* optimal allocations. It also allows the introduction of new deviation measures to be used in portfolio optimization, in order to overcome some of the limits of the classical mean–variance framework. As an example, using asymmetric error measures, it is possible to develop deviation measures that allow one to penalize asymmetrically positive and negative movements, suitable for investors that are averse to losses. In contrast, standard deviation, weighting in the same way positive and negative deviations from the mean, may result in a mis-evaluation of risk in case of assets with asymmetric distributions.

*Corresponding author. Email: sandra.paterlini@unitn.it

Beyond the debate on the characteristics of risk measures, in practical applications the effectiveness of an asset allocation strategy cannot be related exclusively to the theoretical properties of the selected risk or deviation measures, but has to take into account the ex-post results. These are affected by the estimation error and the statistical properties of the data. Therefore the choice of a model for portfolio optimization (and hence of a suitable measure of risk/deviation) should be also guided by its ability to perform well out-of-sample.

Indeed, one of the main challenges of portfolio strategies determined by solving an optimization problem is to exhibit satisfactory out-of-sample performance, especially when estimation error, unrealistic distributional assumptions or mis-specification of the model can lead to poor results. For instance, in the mean–variance framework, Michaud (1989) shows how optimization can act as *error maximization*, and DeMiguel *et al.* (2007) show that the naïve equally weighted $1/n$ portfolio is characterized by a better Sharpe ratio, certainty-equivalent returns and turnover than the mean–variance model and many of its extensions. Pflug *et al.* (2012) demonstrate that the uniform investment strategy is rational in situations where an agent is faced with a sufficiently high degree of model uncertainty in the form of ambiguous loss distributions.

Several authors proposed techniques to improve out-of-sample performance, such as shrinking the covariance matrix (Ledoit and Wolf 2004), using random matrix theory for covariance estimation (Laloux *et al.* 2000), regularize the covariance matrix by imposing a penalization on the elements of its inverse (Goto and Xu 2015, Torri *et al.* 2019) or directly penalizing the portfolio weights in quantile regression (Bonaccolto *et al.* 2018).

The problem of out-of-sample performance is even more pronounced when tail risk measures are considered, as their estimation is typically more challenging than the variance. Salahi *et al.* (2013) propose a robust optimization scheme for the mean-CVaR based on interval and ellipsoidal uncertainty sets, and Ban *et al.* (2016) suggest performance-based regularization of the problem that constrain the solution in a set characterized by lower out-of-sample variability of the objective function.

A common approach to portfolio regularization is the introduction of an ℓ_1 (lasso) penalty† on the portfolio weights (see Tibshirani 1996 for ℓ_1 penalty in general, Fan *et al.* 2012 for the mean–variance portfolio, and Gendreau *et al.* 2019 for *CV aR*). Also the penalized quantile-based asset allocation proposed by Bonaccolto *et al.* (2018) can be redirected to the penalized optimal *CVaR* framework. Beyond the ℓ_1 penalty, common alternatives are the ℓ_2 penalty (ridge), elastic-net (i.e. a weighted average of ℓ_1 and ℓ_2 penalty), and non-convex penalties such as ℓ_q and SCAD (Fastrich *et al.* 2015, Giuzio and Paterlini 2019).

In this paper, we consider asset allocation frameworks that minimize deviation measures. In particular, we introduce the expectile-based risk quadrangle, with the corresponding deviation measure. The new measure built on expectiles is compared to one constructed on quantiles proposed by Bassett *et al.* (2004). In order to improve the out-of-sample performance, especially for application to large portfolios, ridge and elastic-net penalizations are proposed as regularization tools. Performances are then evaluated out-of-sample on several real-world datasets. The goal of the proposed allocations is to better control for tail risk thanks to the asymmetric measures, as well as limiting the effect of estimation error due to the introduction of ridge and elastic-net penalties.

The main contributions of the paper are the following: first, to our knowledge we are the firsts to discuss deviation measures based on the expectile in the risk quadrangle framework of Rockafellar and Uryasev (2013), and to use them for portfolio management. Second, we consider regularized versions of the aforementioned asset allocation strategies based on ridge and elastic-net penalization, showing how such regularization can improve out-of-sample performances.

The paper is organized as follows: in section 2 we introduce two deviation measures and we discuss their characteristics in relationship to the ‘risk quadrangle’ proposed by Rockafellar and Uryasev (2013). In section 3, we discuss the penalized asset allocations and we introduce quadratic programming formulations of the optimization problem; in section 4, we conduct a simulation study to assess the performance of penalized and unpenalized allocations; section 5 consists of an empirical analysis on real data; and section 6 concludes.

2. Optimization of deviation measures

2.1. Deviation measures

A deviation measure describes the dispersion of a random variable around a measure of central tendency. Variance and standard deviation are the most commonly used in the financial literature. Other examples are the mean absolute deviation (MAD) or the semivariance.

As it is well known in statistics, deviation measures can be interpreted as the solution of variational problems. Considering a random variable X, a deviation measure can be expressed as follows:

$$\mathcal{D}_{\zeta}(X) = \min_{\xi} \mathbb{E}[\zeta(X - \xi)] \tag{1}$$

where ζ is a generic error function, $\xi \in \mathbb{R}$ and X is a univariate random variable.

Simply by changing the function $\zeta(X)$, we can derive different deviation measures. The value ξ^* that minimizes the objective function is the centrality measure associated with the deviation measure. As an example, for X being a square-integrable random variable, using the function $\zeta(X) = X^2$ we obtain the variance and $\xi^* = \arg\min_{\xi} \mathbb{E}[(X - \xi)^2]$ is then the mean.

In this work, we focus on two deviation measures derived from the quantile regression and expectile regression, respectively. Compared to the variance, such measures are asymmetric and should therefore be better suited to the application to portfolio optimization in presence of investors’ loss aversion and asymmetry in equity returns, which is a typical stylized fact (Cont 2001). The measures we consider are:

† The ℓ_1 and ℓ_2 penalties are proportional to the 1-norm and 2-norm of the asset weights vector, respectively.

(i) $\mathcal{D}_{\rho_\tau}$—*Quantile-based deviation.* The first measure can be obtained considering the modified Koenker and Bassett error function $\rho_\tau(X)$, used in the estimation of the quantile regression and defined as follows:

$$\rho_\tau(X) := X_+ + (1-\tau)/\tau X_-, \tag{2}$$

where $\tau \in (0,1)$ is a scalar that controls the shape of the function, X is a random variable in L^2, $X_+ = \max(0, X)$ and $X_- = \max(0, -X)$ (for the use in the quantile regression framework, see e.g. Koenker and Bassett 1978). From (1), we have

$$\mathcal{D}_{\rho_\tau}(X) = \min_\xi \mathbb{E}[\rho_\tau(X-\xi)] \tag{3}$$

$$q_\tau(X) = \xi^* = \arg\min_\xi \mathbb{E}[\rho_\tau(X-\xi)] \tag{4}$$

where $q_\tau(X)$ is the τ-quantile of the variable X and $CVaR_{1-\tau}(X) = \mathcal{D}_{\rho_\tau}(X) - \mathbb{E}[X]$.

(ii) $\mathcal{D}_{\eta_\tau}$—*Expectile-based deviation.* We then consider the error function η_τ, used in the estimation of expectile regression (Newey and Powell 1987), defined as follows:

$$\eta_\tau = (X-\xi)_+^2 + (1-\tau)/\tau(X-\xi)_-^2, \tag{5}$$

with $\tau \in (0,1)$ and $X \in L^2$. We obtain the following deviation measure and centrality measure:

$$\mathcal{D}_{\eta_\tau}(X) = \min_\xi \mathbb{E}[\eta_\tau(X-\xi)], \tag{6}$$

$$EVaR_{1-\tau}(X) = -\xi^* = -\arg\min_\xi \mathbb{E}[\eta_\tau(X-\xi)]. \tag{7}$$

Equation (7) corresponds to the expectile, an increasingly popular measure in the financial literature (see, e.g. Bellini and Bignozzi 2015, Bellini and Di Bernardino 2017, Bellini *et al.* 2018). Concerning the notation, we adopt the convention of Bellini and Di Bernardino (2017), calling the risk measure $EVaR_{1-\tau}(X)$ analogously to the most common notations for Value at Risk ($VaR_{1-\tau}(X)$) and Conditional Value at Risk ($CVaR_{1-\tau}(X)$).

2.2. *Deviation measures and optimal portfolios*

Let us consider an $[n \times 1]$ vector of weights $\boldsymbol{w} \in \mathbb{R}^n$, where n is the number of assets, and τ the confidence level with $\tau \in (0,1)$ for the quantiles, superquantiles, and expectiles. A portfolio optimization scheme based on dispersion measures can be formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^n,\xi\in\mathbb{R}} \quad & \mathbb{E}[\zeta(R\boldsymbol{w}-\xi)] \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1 \\ & \mathbb{E}[R\boldsymbol{w}] = \kappa \end{aligned} \tag{8}$$

where $\mathbf{1}$ is a $[n \times 1]$ vector of ones, R is an n-variate random variable denoting stock returns, and κ is a scalar representing a target expected return. The two constraints represent the budget constraints and a target expected return. If we omit the expected return constraint, we obtain instead the global minimum deviation portfolio. The model can then be extended considering additional constraints, such as limits to the turnover, upper and lower bounds, or stochastic dominance constraints. The general model nests many specific ones, including the mean–variance and the mean-MAD (minimum absolute deviation) portfolios.

Concerning the dispersion measures described in section 2.1, Bassett *et al.* (2004) propose an application of the dispersion measure $\mathcal{D}_{\rho_\tau}$, while, to our knowledge, the deviation measure $\mathcal{D}_{\eta_\tau}$ related to the expectile has not been used for portfolio applications so far.†

Note that, due to the shape of the error functions $\rho_\tau(X)$ and $\eta_\tau(X)$, for $\tau < 0.5$ the corresponding deviation measures weight more the negative returns than the positive ones. In comparison, the variance weights equally the positive and negative deviations from the mean.

In the empirical application, we implement asset allocation schemes based on $\mathcal{D}_{\rho_\tau}$ and $\mathcal{D}_{\eta_\tau}$. Additionally, we consider penalized versions of (8) in order to regularize and sparsify the solution, ideally then improving the out-of-sample performances.

2.3. *Relation between deviation and risk measures: does it matter?*

The deviation-return portfolio in (8) is an extension of the mean–variance scheme introduced by Markowitz (1952). In the last decades, however, the literature focused on portfolios constructed on the basis of *risk measures* rather than *deviation measures*. We show here that the deviation-return portfolio in (8) can be interpreted equivalently as a risk-return portfolio.

In particular, the quadrangle is composed by four quantities, related to each other starting from one statistic $\mathcal{S}(X)$: a risk measure $\mathcal{R}(X)$, a deviation measure $\mathcal{D}(X)$, an error function $\mathcal{E}(X) = \mathbb{E}[\zeta(X)]$ and a regret function $\mathcal{V}(X) = \mathbb{E}[\zeta(X) - X]$, where $\zeta(X)$ is a function on $(-\infty, \infty)$.

The deviation measure $\mathcal{D}(X)$ is computed as the minimum of $\mathcal{E}(X-\xi)$, where $\xi \in \mathbb{R}$, and the risk measure $\mathcal{R}(X)$ is the minimum of $\mathcal{V}(X-\xi)$. Moreover, the following relationships hold:‡

$$\mathcal{R}(X) = \mathcal{D}(X) - \mathbb{E}[X], \tag{9}$$

$$\mathcal{V}(X) = \mathcal{E}(X) - \mathbb{E}[X]. \tag{10}$$

Under the quadrangle framework, the optimization problem (8) aims at finding the weights that minimize the deviation measure $\mathcal{D}(R\boldsymbol{w})$ under the budget constraint and a constraint on the expected return of the portfolio.

† Hu and Zheng (2013) propose a version of CAPM that includes view bias, and introduce the *variancile*, a risk measures that corresponds to $\mathcal{D}_{\eta_\tau}$.

‡ A detailed description of the quadrangle framework is outside the scope of this work, and the reader can refer to Rockafellar and Uryasev (2013). Note that the sign convention adopted here is opposite to the one used in Rockafellar and Uryasev (2013), as we model the returns of assets, while they model losses.

Using the quadrangle, we can characterize the relationship between minimal risk and minimal deviation portfolios, and establish the following relationship:

PROPOSITION 2.1 *Any portfolio* $\boldsymbol{w}_d^*$ *that minimizes the deviation measure* $\mathcal{D}(R\boldsymbol{w})$ *under a return constraint* $\mathbb{E}[R\boldsymbol{w}] = \kappa$ *with* $\kappa \in \mathbb{R}$*, computed as*

$$\begin{aligned} \boldsymbol{w}_d^* = \arg \min_{\boldsymbol{w}\in\mathbb{R}^n,\ \xi\in\mathbb{R}} \quad & \mathcal{E}(R\boldsymbol{w}) \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1 \\ & \mathbb{E}[R\boldsymbol{w}] = \kappa, \end{aligned} \tag{11}$$

is equal to the portfolio $\boldsymbol{w}_r^*$ *that minimizes the risk* $\mathcal{R}(R\boldsymbol{w})$ *among all the portfolios with expected return* κ *by solving the following problem:*

$$\begin{aligned} \boldsymbol{w}_r^* = \arg \min_{\boldsymbol{w}\in\mathbb{R}^n,\ \xi\in\mathbb{R}} \quad & \mathcal{V}(R\boldsymbol{w}) \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1 \\ & \mathbb{E}[R\boldsymbol{w}] = \kappa, \end{aligned} \tag{12}$$

Proof According to equation (10), the objective function in (12) can be expressed as $\mathcal{E}(X) - \mathbb{E}[X]$. Since the problems (11) and (12) have the same constraint set, and the value of $\mathbb{E}[X]$ is constant, it follows that the maximizer of (11) is equal to the maximizer of (12). That is, $\boldsymbol{w}_d^* = \boldsymbol{w}_r^*$. ■

REMARK Both the minimum risk and minimum deviation portfolios in presence of the expected return constraint can be represented as special cases of the following problem:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^n,\ \xi\in\mathbb{R}} \quad & \mathbb{E}[\zeta(R\boldsymbol{w} - \xi)] - \delta\mathbb{E}[R\boldsymbol{w}] \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1, \end{aligned} \tag{13}$$

where the constraint on the expected return has been substituted by an additional component of the objective function. For the deviation-return portfolio, δ is the Lagrange multiplier of the expected return constraint, while for the risk-return portfolio δ is equal to one plus the Lagrange multiplier.

REMARK A portfolio that lies on the deviation-return efficient frontier is not necessarily risk-return efficient. An example is the global minimum deviation portfolio. All the efficient risk-return portfolios instead are also deviation-return efficient. As it is shown in figure 1, this is due to the fact that some of the minimum deviation portfolios lie on the lower section of the risk-return curve, and therefore it is possible to find portfolios with the same risk, but with higher expected return.

Focusing on actual asset allocations, the relationship highlighted above shows that the minimum deviation portfolios based on the measure $\mathcal{D}_{\rho_\tau}$ and an expected return constraint, is equivalent to the mean-*CV aR* portfolio with the same expected return constraint.

Figure 1 shows both the risk-return (left) and deviation-return (right) frontiers, highlighting the position of the efficient Global Minimum Risk (GMR, corresponding to $\delta = 1$) and Global Minimum Deviation (GMD, corresponding to $\delta = 0$) portfolios. Notably, as shown in Rockafellar and Uryasev (2013, section 2), the deviation measure $\mathcal{D}_{\rho_\tau}$ is related to the risk measure *CV aR*. It follows therefore that the optimal mean-deviation portfolio constructed using $\mathcal{D}_{\rho_\tau}$ lies on the same frontier of the optimal mean-*CV aR* portfolio.

3. Penalized optimal deviation asset allocations

Portfolio optimization over a large set of asset is typically characterized by unsatisfactory out-of-sample performances, as a consequence of the estimation error of the model parameters and the large dimensionality of the problem (see, e.g. Jagannathan and Ma 2003, Ledoit and Wolf 2004, Fastrich *et al.* 2015). A popular approach aiming at improving out-of-sample performances is to use penalization techniques that sparsify and/or shrink the estimates, leading to more stable solutions.

The use of regularization stems from the statistical literature, starting from the lasso regression proposed by Tibshirani (1996), that regularizes the solution by introducing an ℓ_1 penalty on the asset weights. Such approach has been extended to other penalty functions and regression frameworks. In particular, for quantile regression see Belloni and Chernozhukov (2011) and for expectiles regression (Liao *et al.* 2019).

In the context of portfolio optimization, we add the penalization to the objective function, such that:†

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^n} \quad & \mathcal{D}(R\boldsymbol{w}) + p_\lambda(\boldsymbol{w}) \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1, \\ & \mathbb{E}[R\boldsymbol{w}] = k, \end{aligned} \tag{14}$$

where $\mathcal{D}(X)$ is a risk measure and $p_\lambda(\boldsymbol{w})$ a penalization function applied on portfolio weights. The function p_λ can be specified in several ways: among the most common are the ℓ_1 penalty (lasso) (i.e. the 1-norm of the vector of asset weights), that shrinks some of the weights exactly to zero, and the ℓ_2 penalty (ridge) (i.e. the 2-norm of the vector of the asset weights), that shrinks the vector of portfolio weights towards the equally weighted portfolio.

We adopt here the *elastic-net* penalty (see figure 2) that combines the ℓ_1 (lasso) and ℓ_2 (ridge) penalties. The advantages of this specification are that it induces sparsity in the weights thanks to the corner point (i.e. shrinks some of the weights to zero due to the 1-norm component), and better deals with collinearity than lasso (due to the 2-norm component), that selects arbitrarily one of the collinear variables (Kremer *et al.* 2020). Moreover, in presence of a budget constraint the lasso penalty is not effective once the portfolio

† Alternatively, it is possible to formalize the problem as a regression model (see appendix 1). This approach has been considered by Fan *et al.* (2012) for global minimum variance portfolio, and by Bonaccolto *et al.* (2018) for the quantile-based allocation. The disadvantage is that it requires the selection of a numeraire asset that is not penalized and that does not easily allow to introduce additional constraints to the optimization model.

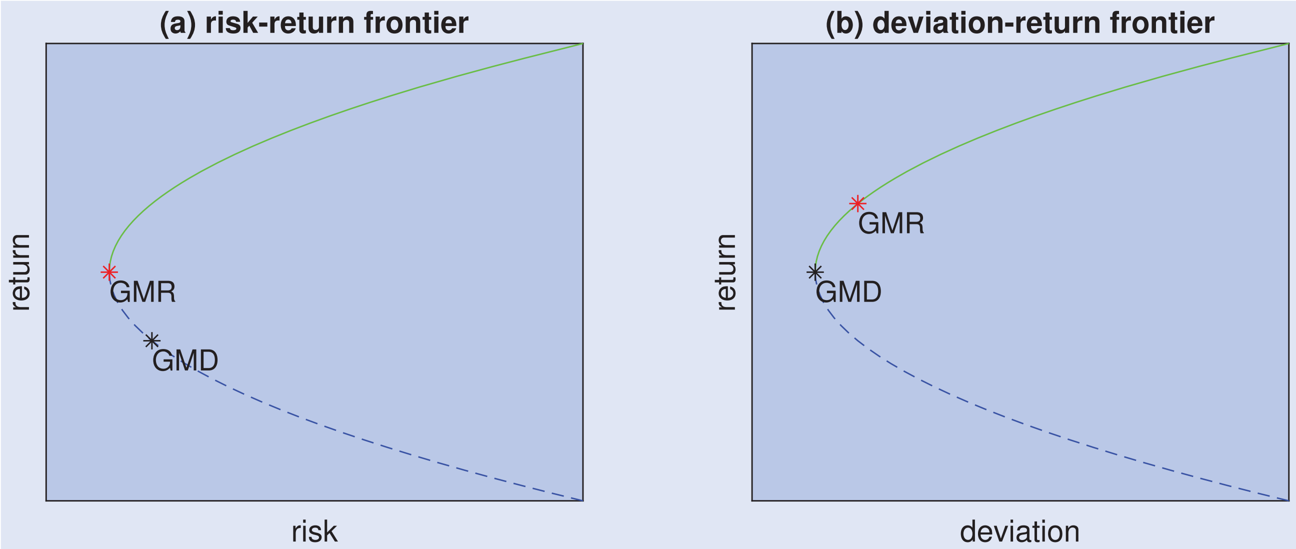

Figure 1. Risk-return and deviation-return frontiers with the global minimum risk (GMR) and global minimum deviation (GMD) portfolios highlighted.

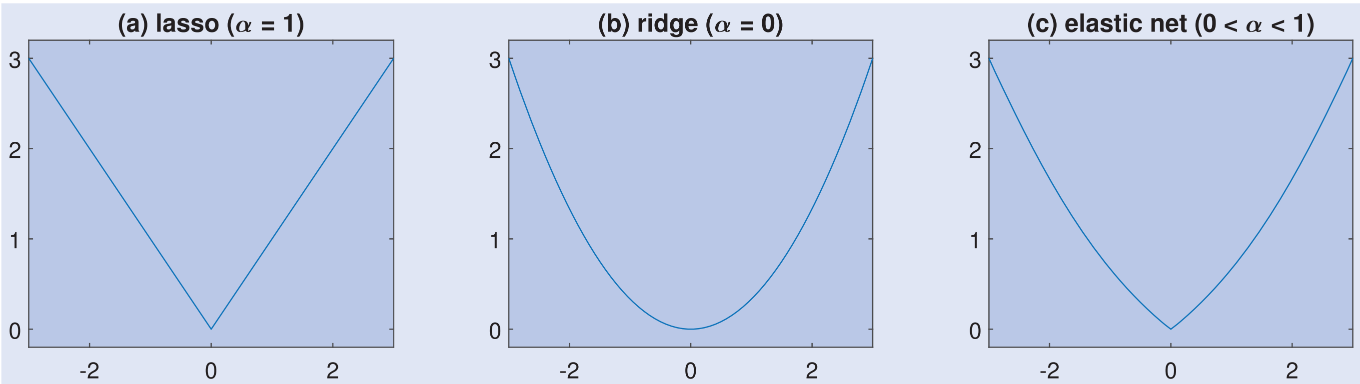

Figure 2. Lasso, ridge and elastic-net penalities.

becomes long only as the 1-norm of the weights is then a constant, while ridge and elastic-net penalties keep regularizing the allocation up to the equally weighted portfolio.

For $\mathcal{D}_{\rho_\tau}$ (quantile-based measure) the elastic-net penalized optimization problem with an expected return constraint is then:

$$\begin{aligned}\mathcal{D}_{\rho_\tau}(R\boldsymbol{w}) = \min_{\boldsymbol{w}\in\mathbb{R}^n,\ \xi\in\mathbb{R}} \quad & \mathbb{E}[\zeta_\tau(R\boldsymbol{w}-\xi)] \\ & + \lambda\left(\alpha||\boldsymbol{w}||_1 + (1-\alpha)||\boldsymbol{w}||_2^2\right) \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1, \\ & \mathbb{E}[R\boldsymbol{w}] = \kappa, \end{aligned} \tag{15}$$

where $\rho_\tau(\cdot)$ is the penalty function associated to the deviation measure $\mathcal{D}_{\rho_\tau}(\cdot)$, λ is a suitably chosen coefficient in $\mathbb{R}^+$, $\alpha \in (0,1)$, $||\cdot||_1$ and $||\cdot||_2$ are the 1- and 2-norm, respectively. The ridge penalized problem can be obtained by setting $\alpha = 0$.

The problem can then be recasted as the following quadratic programming by including t variables γ_i, t variables ϕ_i and n variables ν_j:

$$\begin{aligned}\min_{\boldsymbol{w}\in\mathbb{R}^n;\ \boldsymbol{\nu}\in\mathbb{R}^n_+;\ \boldsymbol{\phi},\boldsymbol{\gamma}\in\mathbb{R}^t_+;\ \xi\in\mathbb{R}} \quad & \mathbb{E}\Bigg\{\tau\sum_{i=1}^t \boldsymbol{\gamma}_i^2 + (1-\tau)\sum_{i=1}^t \boldsymbol{\phi}_i^2 \\ & + \lambda\sum_{j=1}^n(\alpha\boldsymbol{\nu}_j + (1-\alpha)\boldsymbol{\nu}_j^2)\Bigg\} \\ \text{s.t.} \quad & \gamma_i - \phi_i = R_i\boldsymbol{w} - \xi \quad \forall\, i = 1,\ldots,t \\ & -\boldsymbol{\nu}_j \le \boldsymbol{w}_j \le \boldsymbol{\nu}_j \quad \forall\, j = 1,\ldots,n \\ & \boldsymbol{w}'\mathbf{1} = 1. \\ & \left(\sum_{i=1}^t R_i/t\right)\boldsymbol{w} = \kappa. \end{aligned} \tag{16}$$

Similarly, the optimization problem for the elastic-net optimal portfolio built using $\mathcal{D}_{\eta_\tau}$ (i.e. based on expectile), is as follows:

$$\begin{aligned}\min_{\boldsymbol{w}\in\mathbb{R}^n,\ \xi\in\mathbb{R}} \quad & \mathbb{E}[\rho_\tau(R\boldsymbol{w}-\xi)] + \lambda\left(\alpha||\boldsymbol{w}||_1 + (1-\alpha)||\boldsymbol{w}||_2^2\right) \\ \text{s.t.} \quad & \boldsymbol{w}'\mathbf{1} = 1, \\ & \mathbb{E}[R\boldsymbol{w}] = \kappa \end{aligned} \tag{17}$$

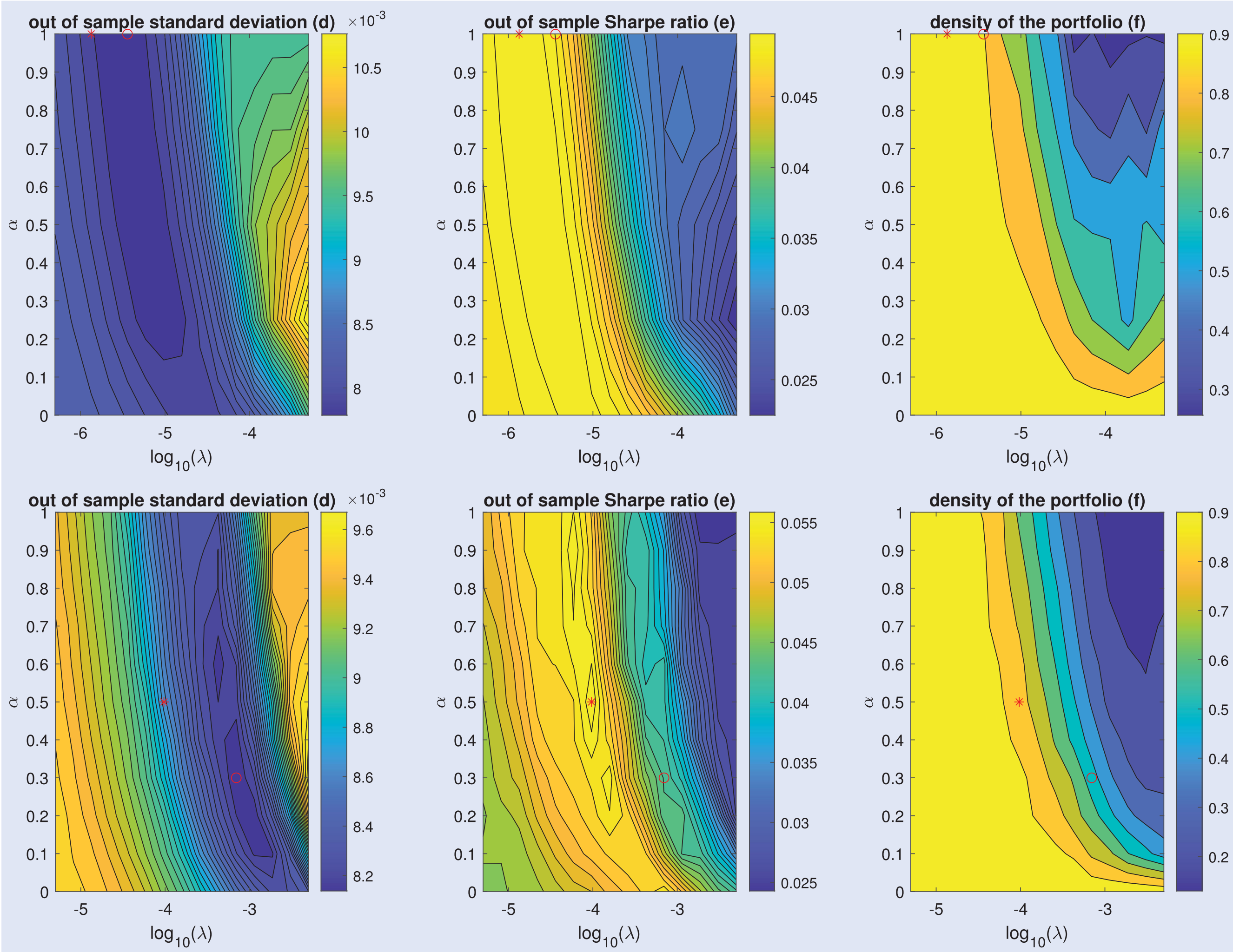

Figure 3. Contour plots (over a grid of values of λ and α) of the out-of-sample standard deviation (first column), Sharpe ratios (second column) and portfolio density (i.e. ratio of weights with non-zero weight, third column) of the minimum $\mathcal{D}_{\eta_\tau}$ portfolios (expectile-based, top panels) and minimum $\mathcal{D}_{\rho_\tau}$ portfolios (quantile-based, bottom panels). The red asterisks denote the portfolios with the highest Sharpe ratios, while the red circles the portfolios with the lowest standard deviation.

and can be formulated as the following quadratic problem:

$$\begin{aligned}
\min_{\boldsymbol{w}\in\mathbb{R}^n;\ \boldsymbol{v}\in\mathbb{R}^n_+;\ \boldsymbol{\phi},\boldsymbol{\gamma}\ \in\mathbb{R}^t_+;\ \xi\in\mathbb{R}} \quad & \mathbb{E}\Bigg\{\tau\sum_{i=1}^{t}\boldsymbol{\gamma}_i+(1-\tau)\sum_{i=1}^{t}\boldsymbol{\phi}_i \\
& +\lambda\sum_{j=1}^{n}(\alpha\boldsymbol{v}_j+(1-\alpha)\boldsymbol{v}_j^2)\Bigg\} \\
\text{s.t.} \quad & \gamma_i-\phi_i=R_i\boldsymbol{w}-\xi \quad \forall\, i=1,\dots,t \\
& -\boldsymbol{v}_j\le\boldsymbol{w}_j\le\boldsymbol{v}_j \quad \forall\, j=1,\dots,n \\
& \boldsymbol{w}'\mathbf{1}=1. \\
& \left(\sum_{i=1}^{t}R_i/t\right)\boldsymbol{w}=\kappa.
\end{aligned} \tag{18}$$

In the rest of the paper, for brevity we refer to problem (15) that minimizes the deviation measure $\mathcal{D}_{\eta_\tau}$ as the *expectile-based portfolio*, and to the problem (17), that minimizes the deviation measure $\mathcal{D}_{\rho_\tau}$ as the *quantile-based portfolio*.

Moreover, as it is well known that the estimation of expected returns is challenging, and portfolio allocation is sensitive to estimation error in asset returns (see, e.g. Michaud 1989, Broadie 1993, Chopra and Ziemba 1993, DeMiguel *et al.* 2007), the empirical analysis in section 5 focuses on asset allocations computed without the expected return constraint $\mathbb{E}[R\boldsymbol{w}]=\kappa$ to avoid issues related to the estimation of expected returns. We refer to these allocations as *global minimum deviation allocations*, opposed to *mean-deviation allocations*.

3.1. *Tuning of regularization parameters*

The model requires the tuning of two parameters: λ, that controls the overall intensity of the penalization, and α, that controls the trade off between the 1-norm and 2-norm in the elastic-net.

A peculiarity of the use of elastic-net in the context of portfolio optimization, opposite to the regression framework, is the behavior of the portfolio in the limit for $\lambda\to\infty$. Differently from the regression framework, in which both the ℓ_1 and

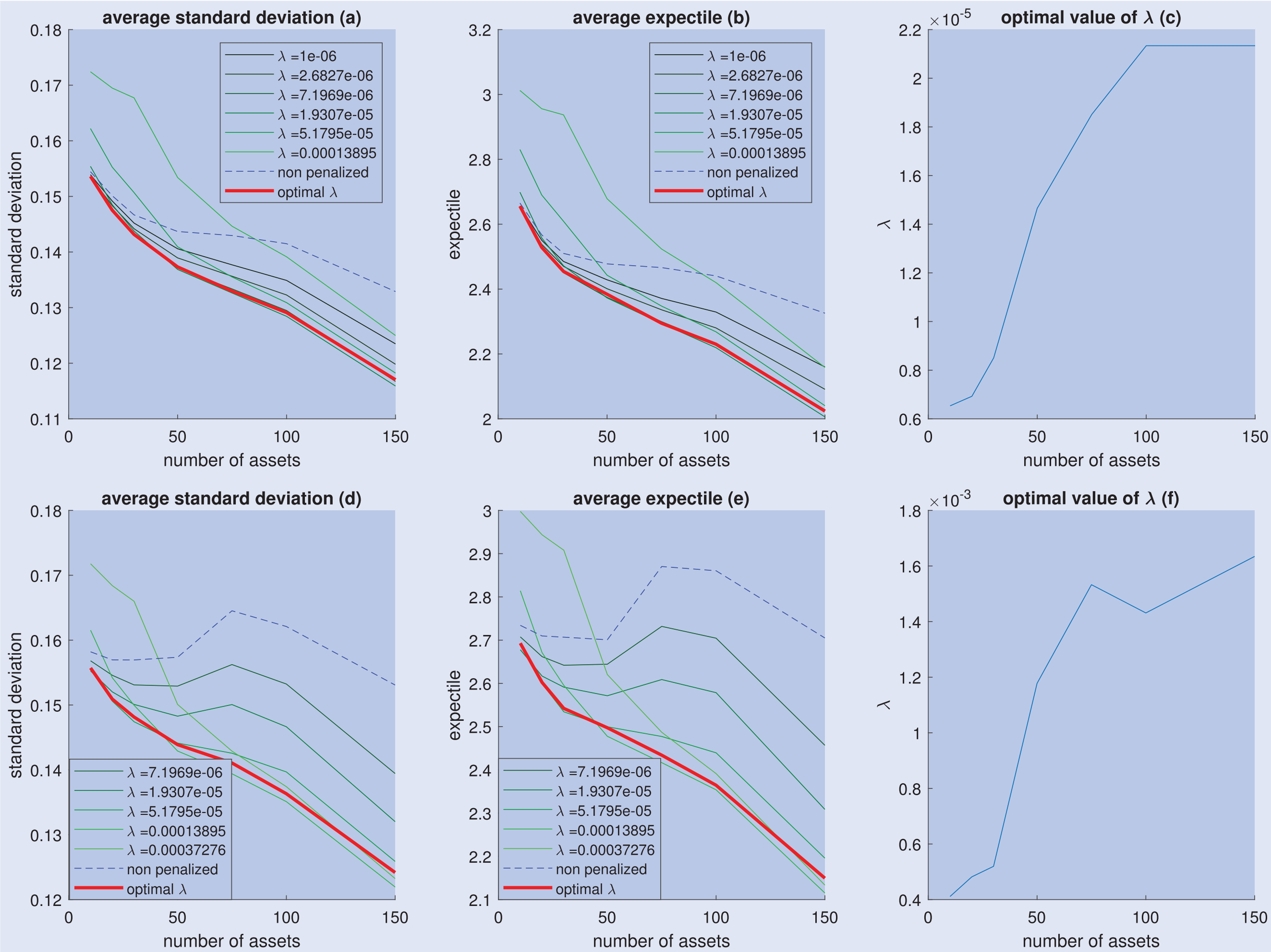

Figure 4. Out-of-sample standard deviation and $EVaR_{95\%}$ for the optimal minimum $\mathcal{D}_{\eta_\tau}$ deviation (expectile-based, top Panels) and optimal minimum $\mathcal{D}_{\rho_\tau}$ deviation (quantile-based, bottom panels). The x-axis represents the number of assets included in the investment universe. Green thin lines denote ridge-penalized portfolios with a fixed value of λ, red bold line denotes the portfolio with λ tuned used cross-validation, and blue, dashed line are the unpenalized portfolios. The plots on the right (c) and (f) represent the optimal value of λ selected using cross-validation.

ℓ_2 penalizations shrink the values of the coefficients towards zero for increasing λs, in portfolio applications the ℓ_1 penalty shrinks the portfolio weights to zero, while the ℓ_2 shrinks them towards $1/n$. The value of α influences therefore the regularization path of the weights: we see first a shrinkage towards zero for increasing values of λ, and then a shrinkage to $1/n$ when the quadratic penalty prevails. This mechanism happens as the ℓ_1 component is the same for all the long-only portfolio. With elastic-net, we have therefore that both the sparsity and the heterogeneity of the weights are controlled by the level of λs, with the balance among these features regulated by the value of α.

Following the approach of Ban *et al.* (2016), we opt for a performance-based tuning obtained using a 10-fold cross-validation technique suitable for time series. In particular, we perform the cross-validation over the last 10 rolling windows as training, using the previous out-of-sample periods (that are then in the past) as test. We then choose the parameter λ that minimizes the standard deviation of the portfolio in the test samples. For α a cross-validation procedure would increase substantially the computational error, while yielding relatively minor benefits, as discussed in section 4. Therefore, in the empirical analysis we fix the value of α to 0.25, that in our study has shown to provide the best empirical performances.

4. Simulation study

The simulation study has the main goal of testing the performance of the regularization procedure and the tuning of the corresponding parameters. The analysis is set up considering a set of random variables jointly distributed as a multivariate t-Student distribution with 5 degrees of freedom. In order to replicate the characteristics of real-world equity markets, we match the first and second moments of the distribution to the returns and sample covariance of the daily returns of a set of 200 assets randomly extracted from the constituents of the Standard and Poor's 500 equity index in the period 01/01/2009–31/12/2018. Unless specified, we do not include any expected return constraint in the optimization, computing therefore minimum deviation portfolios.

Before discussing the tuning of the elastic-net penalty, we assess how the out-of-sample performances are affected by

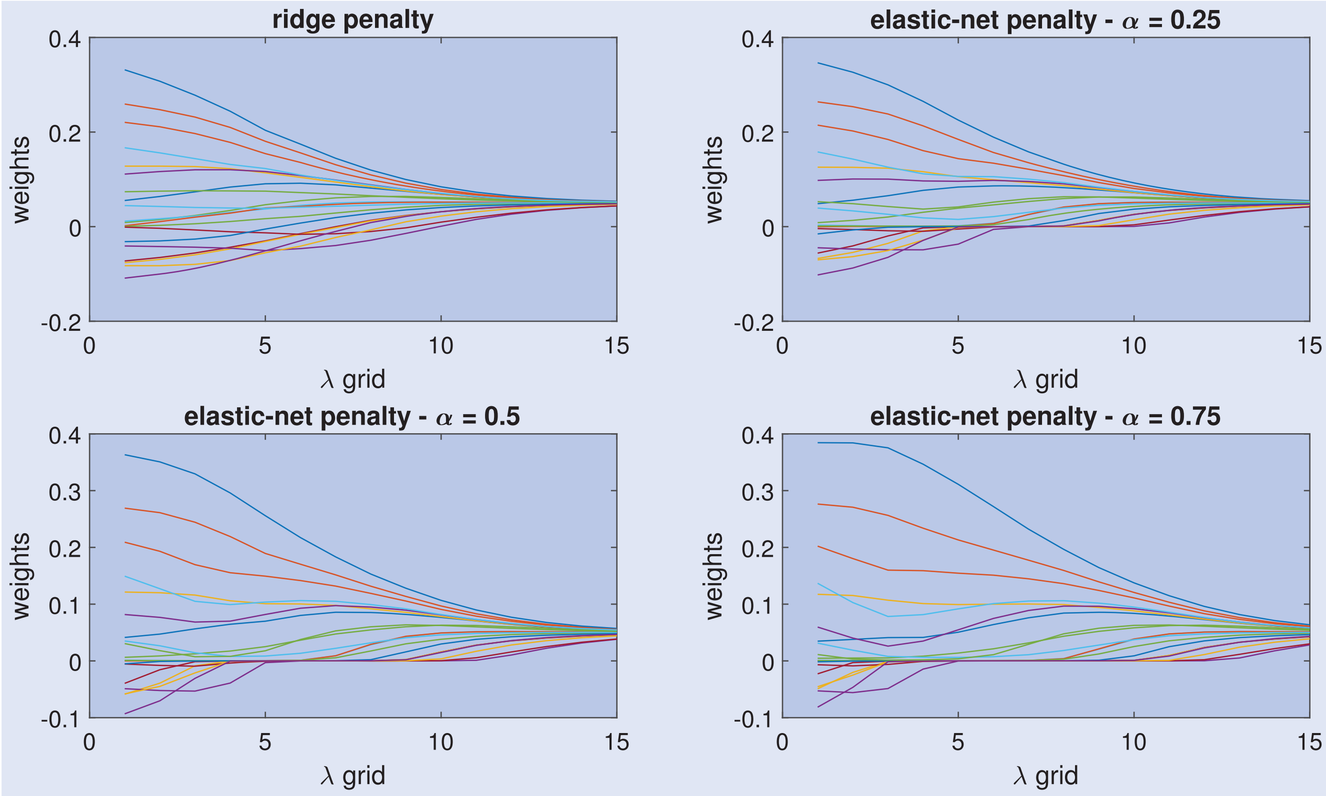

Figure 5. Optimal weights for the minimum $\mathcal{D}_{\eta_\tau}$ portfolio constructed using 20 simulated assets, for different penalization values λ and different values of α. The *ridge* portfolio corresponds to *elastic-net* with $\alpha = 0$.

the choice of the parameters. Figure 3 shows the out-of-sample standard deviation and Sharpe ratio for the optimal minimum $\mathcal{D}_{\eta_\tau}$ portfolios with elastic-net penalization, using different values of λ and α. The portfolios are computed on a subset of 100 assets with 500 observations. The reported results are averaged across 50 simulation runs. We see that the area of the parameter space characterized by the lowest standard deviation and highest Sharpe ratio is a vertical strip. This shifts the attention to the tuning of the λ parameter, as for any given value of α it is possible to identify a portfolio with small risk and high Sharpe ratio.

We then assess the performance of the tuning procedure for λ as outlined in section 3.1. Since the level of regularization is strongly affected by the dimensionality of the data (see, e.g. Fastrich *et al.* 2015), we consider portfolios constructed on investment sets of different sizes, ranging from 10 to 150. The optimal minimum $\mathcal{D}_{\eta_\tau}$ and $\mathcal{D}_{\rho_\tau}$ portfolios have been constructed according to (16) and (18), respectively, on samples of 500 observations. The performances are tested on out-of-sample periods of the same length. The value of τ is set to 0.05, and the results reported below are the average across 50 simulation runs.

Figure 4 shows, for portfolios characterized by different values of λ (i.e. different levels of penalization), the level of the out-of-sample standard deviation, the risk measured as the $EVaR_{0.95\%}$,† and the density of the portfolios (i.e. the ratio of assets with non-zero weights). The panels on the top refer to the minimum $\mathcal{D}_{\eta_\tau}$ (expectile-based) portfolios, while the ones on the bottom to minimum $\mathcal{D}_{\rho_\tau}$ (quantile-based) portfolios. As here we focus on the tuning of λ, we report only the results for the ridge-penalized portfolios, as the results are analogous for elastic-net penalization. In general, we see that the tuning procedure works as expected, selecting portfolios characterized by low deviation and low risk for any size of the investment universe. Portfolios constructed using larger investment universes are typically characterized by a higher value of the optimal λ, denoting the need of relatively higher level of regularization to obtain optimal out-of-sample performances. The results are consistent for both the deviation measures considered in the optimization, although we notice that the distance between the unpenalized and the optimally tuned penalized portfolio diverges quicker for minimum $\mathcal{D}_{\rho_\tau}$ portfolios compared to minimum $\mathcal{D}_{\eta_\tau}$ portfolios when the number of assets n grows. The performances of the optimally penalized portfolios instead are comparable between the two frameworks.

We then test the role of the parameter α that balances the role of ℓ_1 and ℓ_2 penalties. As discussed in section 3, the combination between elastic-net penalization and the budget constraint has an interesting effect on the portfolio weights: at first, an increase in the value of λ, shrinks some of the weights to zero, due to the ℓ_1 component. For larger values of λ the ℓ_2 component becomes dominant and the weights shrink to the equally weighted portfolio (as the value of the ℓ_1 component is the same for all the long-only portfolios). Figure 5 displays this behavior on a portfolio constructed using 20 simulated assets, showing that the ridge penalized (i.e. $\alpha = 0$)

† Using the notation of Bellini and Di Bernardino (2017), $EVaR_{0.95\%}$ is the expectile with $\tau = 0.05$. The results are similar when considering other risk measures such as $CVaR$, VaR. We do not report the results for brevity.

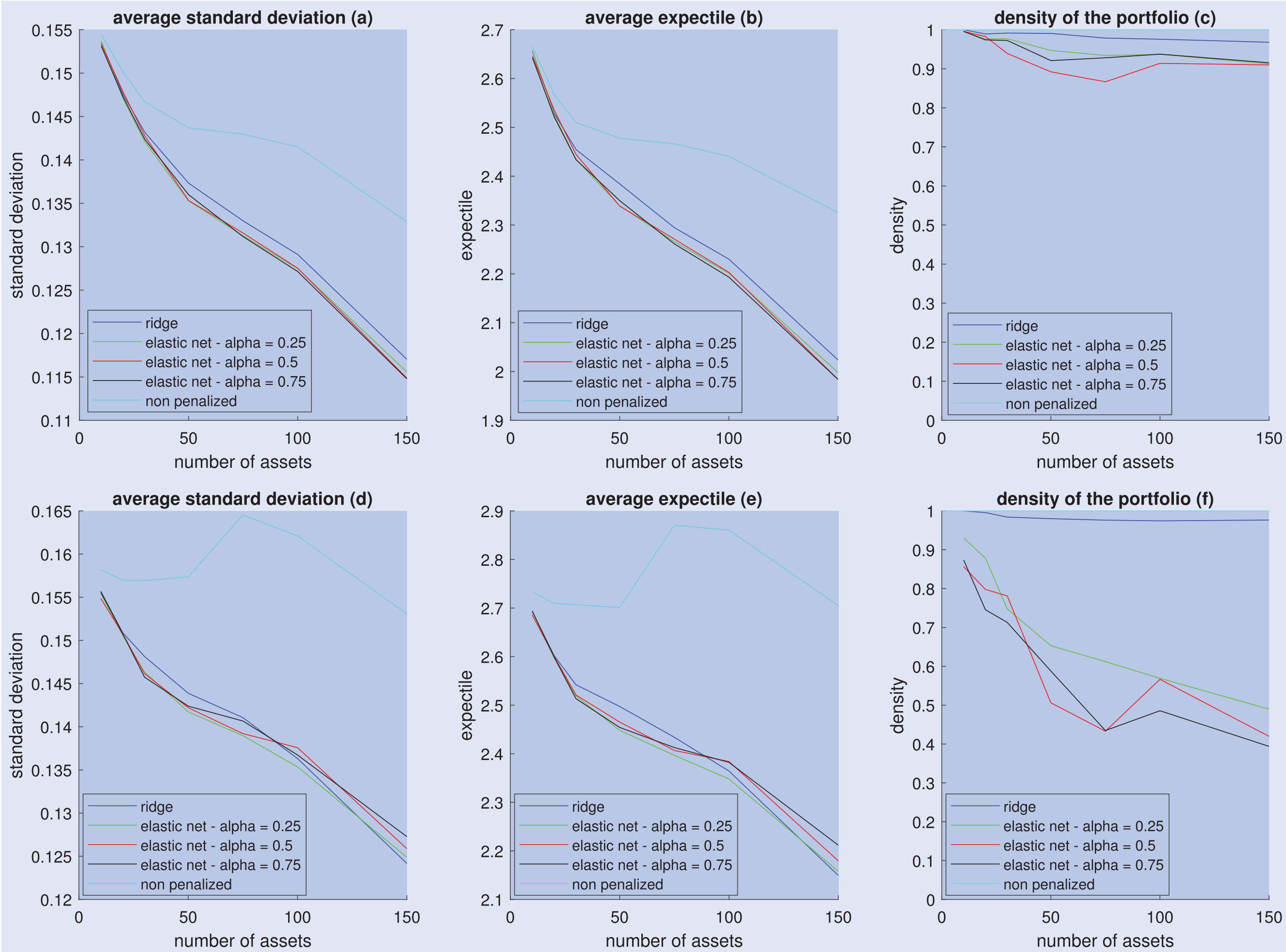

Figure 6. Out-of-sample standard deviation, $EVaR_{95\%}$, and density of the portfolios for the optimal minimum $\mathcal{D}_{\eta_\tau}$ deviation portfolios (top Panels) and optimal minimum $\mathcal{D}_{\rho_\tau}$ portfolios (bottom panels) for different levels of the parameter α. The x-axis represents the number of assets included in the investment universe. λ is tuned using cross-validation. Density is computed as the ratio of assets whose portfolio weights are different from zero.

portfolios converge smoothly to the equally weighted portfolio, while for increasing values of α, an increasing number of weights is set to zero, before converging to the equally weighted portfolio.

Figure 6 shows, for investment universes with different number of assets, the out-of-sample performances of the portfolios (expressed in terms of standard deviation and $EVaR_{0.95\%}$), and the density of the portfolios (i.e. the ratio of assets with non-zero weight), for both the $\mathcal{D}_{\eta_\tau}$ and $\mathcal{D}_{\rho_\tau}$ portfolio allocations. For the $\mathcal{D}_{\eta_\tau}$ case (top Panels), we see that all the values of α lead to approximately the same performances both in terms of risk and standard deviation, and only the portfolios with ridge penalization (i.e. $\alpha = 0$) have slightly worse performances. In the case of $\mathcal{D}_{\rho_\tau}$ (bottom panels), the elastic-net penalized portfolios, with $\alpha = 0.25$ seem to have slightly better performances. Overall all the penalized portfolios have a very similar behavior. Concerning sparsity, in the case of the expectile-based allocations, the penalized portfolios have a high density (above 80%), even for high level of α. In line with the results in figure 3, this could be related to the fact that the optimal values of λ chosen by cross-validation are small enough not to cause sparsity. The quantile-based portfolios instead are more sparse, especially when the investment universe is large.

Finally, we consider the effect of the expected return constraint. Previous literature on penalized portfolio optimization focused on minimum deviation portfolios, due to the problems related to estimation error of the expected returns. In particular, we assess the relative performance of the penalized portfolios in relationship with the amount of penalization (regulated by the parameter λ) and the balance between ℓ_1 and ℓ_2 penalties (regulated by the parameter α). Figure 7 shows the level of out-of-sample standard deviation and expectile, as well as the density of the portfolio, for a grid of target expected returns. All the portfolios are computed on a set of 100 assets. We see ridge and elastic-net are effective in regularizing the portfolios for all the level of target expected return considered. Focusing on the $\mathcal{D}_{\eta_\tau}$ portfolios (figure 7, top panels), the elastic-net portfolio with $\alpha = 0.25$ shows the best performances overall, with lower risk and deviations compared to the other approaches (figure 7, panels a,b), and Sharpe ratio aligned to the one of ridge penalty (figure 7, panel c). Concerning the $\mathcal{D}_{\rho_\tau}$ portfolios (figure 7, bottom panels), the performances of the penalized portfolios are aligned

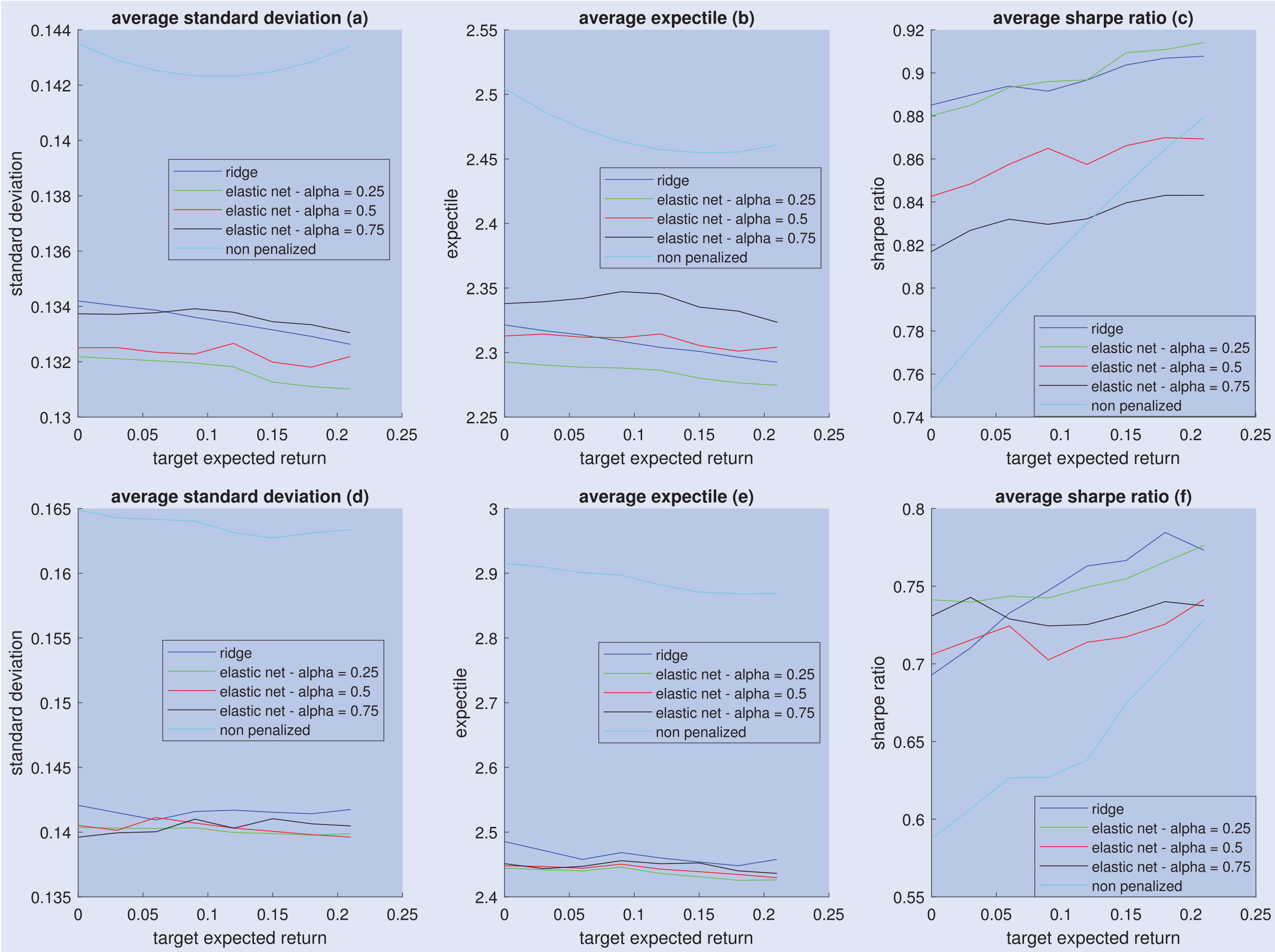

Figure 7. Out-of-sample standard deviation, $EVaR_{95\%}$, and Sharpe ratio for the optimal minimum $\mathcal{D}_{\eta_\tau}$ portfolios (top panels) and optimal minimum $\mathcal{D}_{\rho_\tau}$ portfolios (bottom panels), with an expected return constraint (x axis). The parameter λ is tuned using cross-validation. density of the portfolios for the optimal minimum $\mathcal{D}_{\eta_\tau}$ deviation portfolios (top panels) and optimal minimum $\mathcal{D}_{\rho_\tau}$ portfolios (bottom panels), for different number of assets (x-axis), and levels of the parameter α.

in terms of risk and deviation (figure 7, panels d and e), while, concerning the Sharpe ratio, ridge and elastic-net with $\alpha = 0.25$ outperform the portfolios with higher α.

Overall, the simulation study shows that the penalization techniques used here can greatly improve the out-of-sample performance of the portfolios that minimize $\mathcal{D}_{\eta_\tau}$ and $\mathcal{D}_{\rho_\tau}$, especially in the case of large portfolios. The study also shows that the tuning procedure proposed for λ is effective in minimizing out-of-sample risk and deviation. Concerning the parameter α, that balances the ℓ_1 and ℓ_2 penalty, the results are not conclusive, resulting in similar performances for different levels of the parameter.

5. Empirical analysis

5.1. Data and set up

We test empirically the performances of the optimal asset allocations proposed on a broad dataset composed by 448 constituents of the Standard and Poor's equity index (SP500), with returns available for the entire studied period. We consider a rolling windows scheme with approximately 2 years of daily data (500 observations) as calibration period for each window, and a holding investment period of approximately 1 month of daily data (20 observations). The main analysis focuses on the period 01/02/2008–01/02/2019 (out-of-sample period: 01/02/2010–01/02/2019). In order to test our investment strategy in distressed markets, we also consider two additional periods characterized by high turmoil: the first corresponds to the global financial crisis (out-of-sample period: 01/01/2007–31/12/2008), the latter to the first part of 2020, characterized by the COVID pandemic (out-of-sample period: 01/01/2020–12/06/2020).

Table 1 summarizes the characteristics of the dataset, referring to the out-of-sample periods considered in the analysis. We see that the distribution of returns in the period with higher turmoil is characterized by higher standard deviation, lower expected returns, and higher correlations. The skewness and kurtosis in these periods are smaller compared to the entire period.

We consider the following portfolio allocations:†

† For brevity, we do not include the elastic-net penalized global minimum variance portfolio, as the performances are aligned to the ones of by the ridge-penalized minimum variance portfolio (MV-R).

Table 1. Summary statistics for the datasets used in the analysis. Annualized expected returns, annualized standard deviation, skewness and kurtosis are the average over all the assets in the dataset (for each indicator standard deviations are reported in brackets).

dataset	exp return	std dev	skewness	kurtosis	correlations
SP500 - 2010–2019	10.8%	26.3%	– 36.2%	12.0	37.2%
	(7.25%)	(7.02%)	(64.0%)	(10.8)	(11.5%)
SP500 - 2007–2008	7.75%	47.6%	33.5%	10.7	46.1%
	(13.7%)	(18.6%)	(58.8%)	(6.04)	(10.9%)
SP500 - 2020	– 7.43%	67.7%	2.21%	6.65	60.7%
	(40.9%)	(21.0%)	(52.3%)	(2.45)	(15.3%)

- *MQ—Minimum$\mathcal{D}_{\rho_\tau}$ portfolio (quantile-based)*. Asset allocation that minimizes the deviation measure $\mathcal{D}_{\rho_\tau}$.
- *MQ-R—Minimum$\mathcal{D}_{\rho_\tau}$ portfolio (quantile-based)—ridge penality*. Asset allocation that minimizes the deviation measure $\mathcal{D}_{\rho_\tau}$, with ridge penalization.
- *MQ-EN—Minimum$\mathcal{D}_{\rho_\tau}$ portfolio (quantile-based)—elastic-net penality*. Asset allocation that minimizes the deviation measure $\mathcal{D}_{\rho_\tau}$, with elastic-net penalization.
- *ME—Minimum$\mathcal{D}_{\eta_\tau}$ portfolio (expectile-based)*. Asset allocation that minimizes the deviation measure $\mathcal{D}_{\eta_\tau}$.
- *ME-R—Minimum$\mathcal{D}_{\eta_\tau}$ portfolio (expectile-based)—ridge penality*. Asset allocation that minimizes the deviation measure $\mathcal{D}_{\eta_\tau}$, with ridge penalization.
- *ME-EN—Minimum$\mathcal{D}_{\eta_\tau}$ portfolio (expectile-based)—elastic-net penality*. Asset allocation that minimizes the deviation measure $\mathcal{D}_{\eta_\tau}$, with elastic-net penalization.
- *MV—Minimum variance portfolio*. Markowitz minimum variance portfolio.
- *MV-R –Minimum variance portfolio—ridge penality*. Markowitz minimum variance portfolio with ridge penalization.
- *EW* —Equally weighted portfolio.

For robustness, for each of these allocations we consider three versions of the optimization problems:

- *GMD—Global Minimum Deviation portfolios*. Characterized by no expected returns constraint.
- *MD—Mean-Deviation portfolios*. Optimized considering an expected return constraint with κ equal to the expected return of the equally weighted portfolio of the stocks in the investment universe.
- *GMD-LO—Global Minimum Deviation portfolios—long only*. Optimized with a non-negativity constraint on the portfolio weights.

The parameter λ in the ridge and elastic-net portfolios are calibrated using cross-validation as described in section 3.1, while the parameter α in the elastic-net model is set to 0.25. For each of the asset allocations, we compute a set of out-of-sample risk and return measures, as well as portfolio statistics (for a description of the measures, see appendix 1).

Table 2. Risk and performance measures for different asset allocation strategies. All the values are computed out-of-sample in a rolling windows scheme for the period 01/02/2010–01/02/2019. Dataset: SP500. τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. Panel A reports the results for the global minimum deviation portfolios (no constraints on the expected returns), panel B for the portfolios with an expected return constraint, panel C for the global minimum deviation portfolio with non-negativity constraint. The best and second best results for each measures are highlighted in bold and italics, respectively. The performance measures are described in Appendix 1.

	EVaR (%)	CVaR (%)	Std (%)	$\sqrt{\mathcal{D}_{\eta\tau}}$ (%)	Sharpe	max DD (%)
Panel A—global minimum deviation portfolios (GMD)—SP500 – 2010–19						
MQ	1.29	2.74	24.3	40.5	0.256	65.9
MQ-R	0.47	1.07	9.36	16.8	*0.975*	**11.0**
MQ-EN	0.50	1.12	9.58	17.1	0.770	12.4
ME	1.07	2.31	20.6	34.0	0.405	55.1
ME-R	*0.45*	**1.01**	*8.94*	**16.0**	0.973	*11.2*
ME-EN	0.46	1.04	9.08	16.3	0.952	11.3
MV	1.05	2.24	20.2	33.0	0.438	53.5
MV-R	**0.45**	*1.02*	**8.93**	*16.0*	**1.030**	12.1
Panel B—mean deviation portfolios (MD)—SP500—2010–19						
MQ	1.29	2.77	24.4	40.8	0.328	66.0
MQ-R	0.49	1.10	9.62	17.1	**0.957**	12.3
MQ-EN	0.50	1.13	9.84	17.5	0.668	15.0
ME	1.06	2.27	20.3	33.6	0.456	56.9
ME-R	**0.46**	**1.04**	*9.19*	**16.3**	*0.947*	*11.4*
ME-EN	0.48	1.07	9.38	16.6	0.794	**11.3**
MV	1.04	2.23	20.0	33.0	0.482	54.9
MV-R	*0.47*	*1.05*	**9.12**	*16.3*	0.906	13.0
Panel C—global minimum deviation long only portfolios (GMD-LO)—SP500—2010–19						
MQ	0.55	*1.22*	10.5	*18.5*	0.648	14.2
MQ-R	0.55	1.24	10.6	18.8	0.677	14.9
MQ-EN	*0.55*	1.24	*10.5*	18.8	0.659	14.7
ME	0.57	1.28	10.9	19.3	0.726	13.7
ME-R	0.57	1.29	11.0	19.6	0.731	*13.5*
ME-EN	0.58	1.29	11.0	19.6	0.774	14.1
MV	**0.53**	**1.18**	**10.2**	**18.1**	**0.796**	**11.8**
MV-R	0.57	1.28	10.9	19.5	*0.791*	14.5
EW	0.85	1.92	15.9	28.9	0.656	24.4

5.2. *Empirical results*

Table 2 reports a set of risk and performance indicators for the different asset allocations considered in the analysis for the period 01/02/2010–01/02/2019, highlighting for

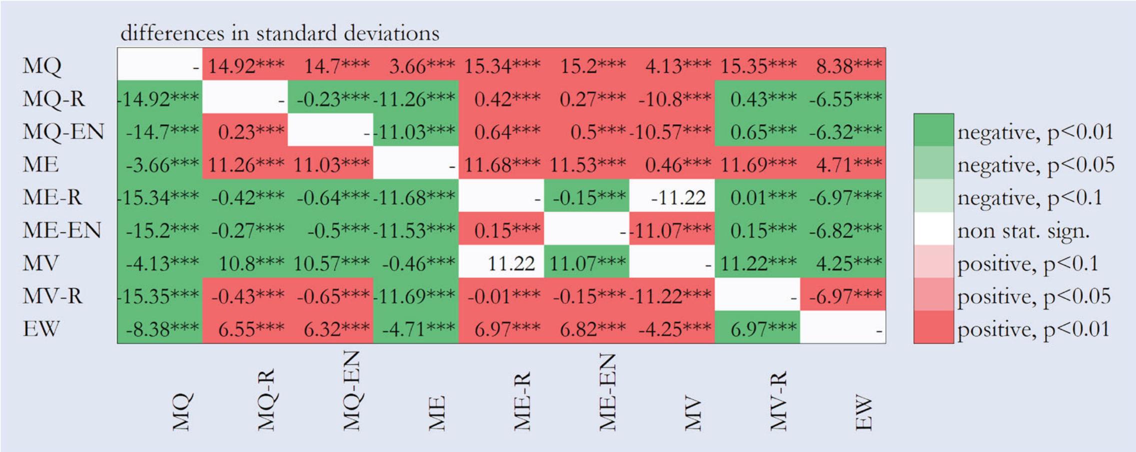

Figure 8. Differences in out-of-sample annualized standard deviation and statistical significance estimated using the Ledoit and Wolf bootstrap procedure (Ledoit and Wolf 2011). *, **, and *** denote significance at the 10%, 5%, and 1%, respectively.

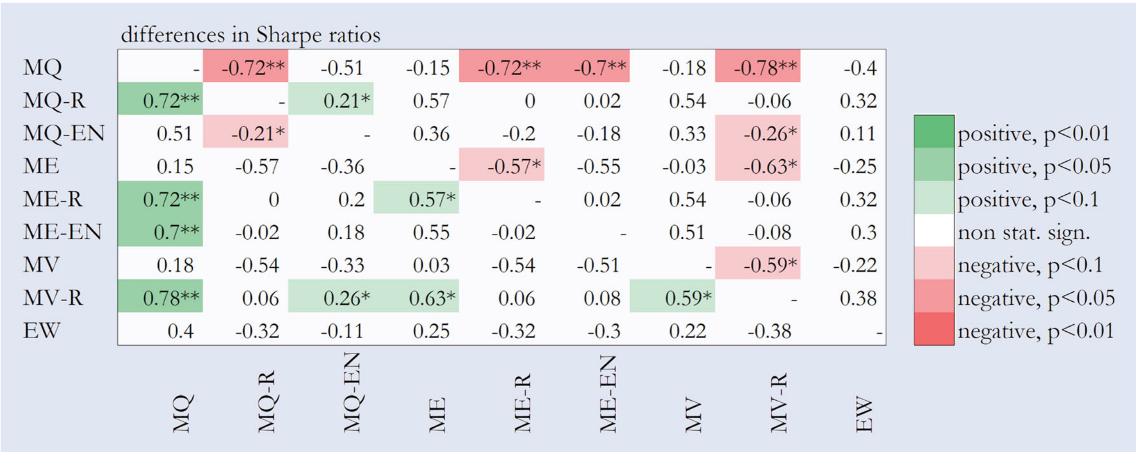

Figure 9. Differences in out-of-sample annualized Sharpe ratios and statistical significance estimated using the Ledoit and Wolf bootstrap procedure (Ledoit and Wolf 2008). *, **, and *** denote significance at the 10%, 5%, and 1%, respectively.

each measure in bold the best performing portfolio, and in italics the second best. We immediately see that the regularization techniques, especially ridge, are effective in reducing the risk of the global minimum deviation (Panel A) and mean-deviation (Panel B) portfolios. As an example, looking at Panel A (Global minimum deviation portfolios), the daily $EVaR_{0.1}$ of the ME portfolio for the SP500 is equal to 1.07%, while the $EVaR_{0.1}$ of the ridge-penalized portfolio (ME-R) is less than half, at 0.45%. The introduction of penalization instead is not beneficial for the long-only portfolios (panel C), probably due to the fact that non-negativity constraints already regularize portfolios weights and narrow the optimization search space. Comparing expectile-based and quantile-based allocations, the formers are more effective in reducing tail risk and deviation compared to the latters in absence of non-negativity constraints (panels A and B), while the ranking in terms of Sharpe ratio and maximum drawdown is less clear. The quantile-based allocations are instead less risky compared to expectile-based in the case of long only portfolios (panel C). More in detail, looking at panel A we see that ME-R and MV-R are the portfolios with the smallest risk, MV-R and MQ-R the ones with the highest Sharpe ratio, and MQ-R and ME-R are the ones with the smallest maximum drawdown. The ranking of the portfolio is similar for Panel B, while in Panel C the unpenalized Markowitz minimum variance portfolios outperform the others in all the considered measures. In general, we see that the Markowitz mean–variance portfolio shows competitive performances in terms of smaller risk in all the frameworks considered. There are two possible explanations for these results. The first is that the estimation of tail measures is indeed typically more challenging than the estimation of covariances, leading to smaller errors (see, e.g. Sarykalin *et al.* 2008, Gotoh and Shinozaki 2010, Danielsson and Zhou 2016). More robust estimation techniques and simulation schemes may provide better performances compared to the use of historical data. We leave such analysis for future research. The second is related to the high consistency of the portfolio ranking according to deviation and risk measures: since portfolio with low standard deviation are also characterized by low risk and deviation measured according to other indicators, the low out-of-sample standard deviation of the Markowitz portfolio is associated with low tail risk. Applications on financial data characterized by highly non-elliptical returns may show a different behavior, highlighting more the benefits of optimization focused on the tails.

We tested the statistical significance of the results by comparing the standard deviations and Sharpe ratios of the proposed asset allocations using the robust tests introduced by Ledoit and Wolf (2011) and Ledoit and Wolf (2008), respectively.† Figures 8 and 9 report the results for the bilateral tests for the GMD portfolios, showing the differences in annualized standard deviations and Sharpe ratios, and the significance according to the bootstrap procedure.‡ We see that the differences in standard deviations between penalized and unpenalized portfolios are in most of the cases significant, confirming the good empirical properties of both the elastic-net and ridge regularization. The data also highlights an advantage of the ridge penalized portfolio, compared to the elastic-net. This result, that differs from the finding of the simulation study, may be caused by the greater unpredictability of the actual data compared to simulated ones. Concerning the Sharpe ratios, with only a few exceptions the differences are not significant, not allowing therefore to draw statistically grounded conclusions when comparing the risk-return efficiency of the portfolios. Finally, regarding the comparison with the equally weighted portfolios, differently from the findings of the literature (see, e.g. DeMiguel *et al.* 2007), our data show that such strategy has statistically significantly higher variance compared to almost all the other portfolios (with the exception of the unpenalized optimal portfolios estimated on the SP500 dataset) and no statistically significantly higher Sharpe ratios.

Table 3 reports portfolio statistics. We see from panels A and B (global minimum deviation and mean-deviation portfolios, respectively), that the penalized strategies greatly reduce turnover, gross exposure and maximum individual exposure, allowing to better control for transaction costs. Concerning portfolio density (i.e. the percentage of assets with non-zero weight), as expected, elastic-net induces more sparsity than ridge penalization.§ Moreover, the quantile-based allocations are sparser than expectile-based ones. Concerning portfolios with non-negativity constraints (panel B), we observe that also in this case the regularization helps in reducing the turnover and the concentration of the portfolios. Differently from the previous cases, the unpenalized portfolios have a smaller density compared to both the elastic-net and ridge penalized. In this case indeed the penalization regularizes the solution by pushing the weights towards the equally weighted portfolio.

Finally, we study the role of the penalization by analyzing the evolution over time of the tuning parameter λ for both ridge and elastic-net penalization in relation to market characteristics (figure 10). We notice that the values of the optimal λs appear to be related to the characteristics of the market, moving together with the volatility, skewness, kurtosis of the market index S&P 500, and correlations of the assets in the market, especially for the ridge penalization. This means that in periods of market distress the cross-validation procedure tends to increase the regularization, leading to asset allocations less sensitive to estimation error and more suitable for periods of high market uncertainty.

Table 3. Portfolio statistics for different asset allocation strategies. All the values are computed out-of-sample in a rolling windows scheme for the period 01/02/2010–01/02/2019. Dataset: SP500. τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. Panel A reports the results for the global minimum deviation portfolios (no constraints on the expected returns), panel B for the portfolios with an expected return constraint, panel C for the global minimum deviation portfolio with non-negativity constraint. The portfolio measures are described in appendix 1.

	Turnover (%)	Gross (%)	Density (%)	Max exp (%)
Panel A—global minimum deviation portfolios (GMD)—SP500—2010–19				
MQ	2650.0	2760.0	98.8	37.3
MQ-R	116.0	362.0	93.1	3.3
MQ-EN	123.0	224.0	32.4	7.5
ME	1760.0	2280.0	98.5	30.3
ME-R	85.2	343.0	91.8	2.8
ME-EN	70.4	271.0	82.2	4.3
MV	1690.0	2230.0	98.5	29.4
MV-R	71.0	316.0	91.4	2.4
Panel B—mean-deviation portfolios (MD)—SP500—2010–19				
MQ	2640.0	2750.0	98.8	37.2
MQ-R	170.0	347.0	67.3	5.4
MQ-EN	138.0	245.0	31.2	8.7
ME	1700.0	2260.0	98.6	29.9
ME-R	113.0	290.0	70.3	5.2
ME-EN	82.2	206.0	54.4	6.9
MV	1670.0	2230.0	98.6	29.5
MV-R	77.3	245.0	67.2	4.0
Panel C—global minimum deviation portfolios long only (GMD-LO)—SP500—2010–19				
MQ	45.8	100.0	5.11	18.7
MQ-R	38.5	100.0	15.1	7.0
MQ-EN	36.2	100.0	14.8	7.1
ME	26.7	100.0	37.5	14.0
ME-R	17.0	100.0	41.7	5.6
ME-EN	18.7	100.0	43.5	7.9
MV	26.6	100.0	6.78	16.0
MV-R	16.6	100.0	32.9	4.2

5.2.1. Portfolio performances in crisis periods. Tables 4 and 5 report the out-of-sample performances and portfolios statistics for the periods 2007–08 (panel A, characterized by the global financial crisis) and first half of 2020 (panel B, COVID pandemic). Overall, the results confirm the ones obtained in the period 2010–19, highlighting even more the role of ridge and elastic-net penalization for the regularization of out-of-sample performance and portfolio exposures. More in detail, in the period 2007–08 (panel A), the MV-R portfolios are the ones characterized by smaller risk, while MQ-EN and ME-EN show the lowest maximum drawdown. Concerning the Sharpe ratios, all the portfolios, with the exception of the equally weighted, show negative values, although quantile- and expectile-based asset allocations perform better than mean–variance portfolios according to such indicator.

† For the implementation we used the Matlab code available on the website of Michael Wolf: https://www.econ.uzh.ch/en/people/faculty/wolf/publications.html. The bootstrap block size is estimated using the data-driven methodology suggested by the authors.
‡ The results for the mean deviation (MD) and global minimum deviation long only (GMD-LO) are not reported for brevity, and are available upon request.
§ Note that we considered a minimum threshold for the weights, therefore also the unpenalized and ridge penalized portfolio do not result perfectly dense.

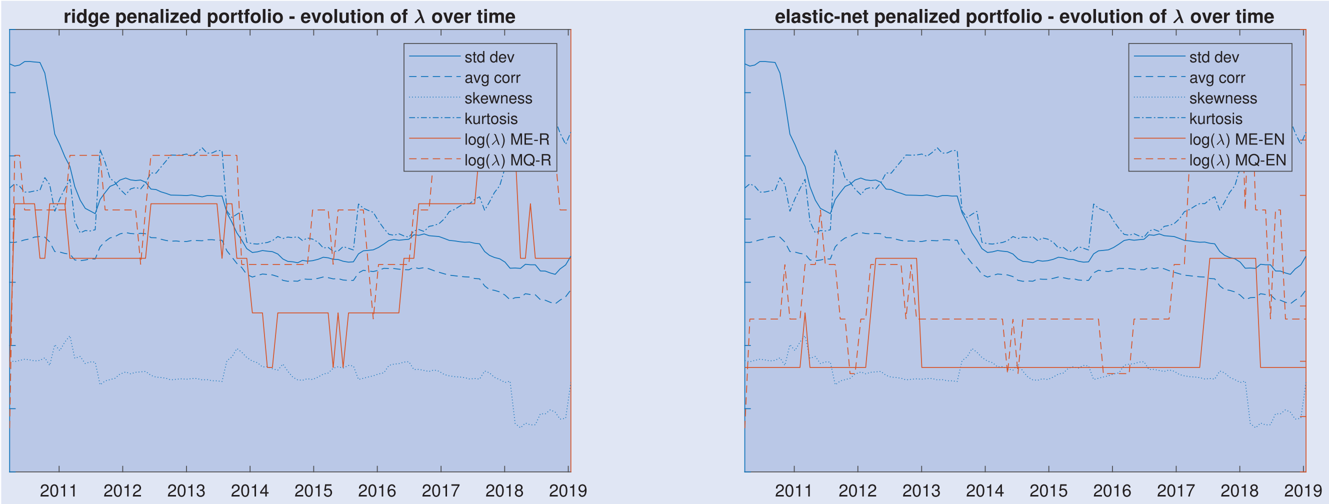

Figure 10. Evolution over time of the optimal value of λ for the ridge penalization (left panel) and elastic net penalization (right panel). The graph also shows the standard deviation, skewness and kurtosis of the S&P 500 index, as well as the average correlation among equity returns. λs are plotted on a log scale, and all the series have been rescaled for better readability.

Table 4. Risk and performance measures for different asset allocation strategies. All the values are computed out-of-sample in a rolling windows scheme for periods 01/01/2007–31/12/2008 (global financial crisis) and 01/01/2020–12/06/2020 (COVID pandemic). Dataset: SP500. τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. Panel A reports the results for the global minimum deviation portfolios (no constraints on the expected returns), panel B for the portfolios with an expected return constraint, panel C for the global minimum deviation portfolio with non-negativity constraint. The best and second best results for each measures are highlighted in bold and italics, respectively. The performance measures are described in appendix 1.

	EVaR (%)	CVaR (%)	Std (%)	$\sqrt{\mathcal{D}_{\eta_\tau}}$ (%)	Sharpe	max DD (%)
Panel A—global minimum deviation portfolio (GMD)—SP500—2007–2008						
MQ	2.18	4.79	40.1	69.3	– *0.07*	56.4
MQ-R	0.82	1.83	15.5	28.8	– 0.17	35.2
MQ-EN	0.82	1.83	15.7	28.9	– 0.08	*31.8*
ME	1.76	3.83	31.9	55.7	– 0.27	55.4
ME-R	0.82	*1.81*	*15.1*	*28.1*	– 0.32	37.7
ME-EN	*0.81*	1.79	15.3	28.5	– 0.13	**31.5**
MV	1.74	3.78	31.1	54.7	– 0.29	55.8
MV-R	**0.78**	**1.72**	**14.3**	**27.4**	– 0.47	40.6
EW	1.73	3.84	32.2	57.3	**0.21**	54.2
Panel B—global minimum deviation portfolio (GMD)—SP500—2020						
MQ	10.90	22.5	206.0	325.0	– 1.14	86.6
MQ-R	*2.35*	**5.16**	38.2	72.2	– 0.44	31.3
MQ-EN	2.50	5.57	42.6	77.2	– **0.28**	**31.1**
ME	11.90	25.2	279.0	402.0	0.71	92.5
ME-R	**2.34**	*5.17*	**38.4**	**73.3**	– 0.47	*31.3*
ME-EN	2.39	5.28	39.6	*74.5*	– *0.41*	31.4
MV	18.70	40.1	328.0	603.0	– 1.00	99.6
MV-R	2.40	5.29	*38.4*	74.7	– 0.59	32.2
EW	3.02	6.46	53.7	94.6	– 0.58	39.5

If we focus on the first quarter of 2020, characterized by the COVID pandemic (panel B), we observe a much higher levels of risk, given the intensity of the crisis. The ME-R portfolio is the one that better controls the risk. Due to the diffused negative returns across all the stocks, the Sharpe ratios are negative for all the portfolios, including the equally weighted. Still, quantile- and expectile-based allocations have less negative Sharpe ratios and lower maximum drawdown compared to the minimum variance and equally weighted portfolios, highlighting the advantages of controlling asymmetric deviation measures in the investment process in periods of market downturns.

Table 5. Risk and performance measures for different asset allocation strategies. All the values are computed out-of-sample in a rolling windows scheme for periods 01/01/2007–31/12/2008 (global financial crisis) and 01/01/2020–12/06/2020 (COVID pandemic). Dataset: SP500. τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. Panel A reports the results for the global minimum deviation portfolios (no constraints on the expected returns), panel B for the portfolios with an expected return constraint, panel C for the global minimum deviation portfolio with non-negativity constraint. The performance measures are described in appendix 1.

	Turnover (%)	Gross (%)	Density (%)	max exp (%)
Panel A—global minimum deviation portfolio (GMD)—SP500—2007–2008				
MQ	2310.0	2350.0	98.5	31.4
MQ-R	143.0	404.0	93.7	4.5
MQ-EN	151.0	263.0	38.7	11.1
ME	1570.0	1930.0	98.0	25.9
ME-R	103.0	369.0	92.9	3.82
ME-EN	50.5	188.0	64.7	12.3
MV	1530.0	1910.0	98.1	25.5
MV-R	82.2	362.0	92.5	3.72
Panel B—global minimum deviation portfolio (GMD)—SP500—2020				
MQ	13200	11100	97.4	123.0
MQ-R	463.0	644.0	95.1	6.64
MQ-EN	302.0	309.0	32.3	10.4
ME	272	20900	99.7	214.0
ME-R	258.0	425.0	92.2	4.44
ME-EN	212.0	271.0	74.7	9.31
MV	30300	18600	99.8	303.0
MV-R	212.0	426.0	93.2	4.39

6. Conclusion

In this work, we study portfolio allocation strategies based on the minimization of two deviation measures: the first related to quantiles and the second related to expectiles. We discuss their properties in the risk quadrangle framework proposed by Rockafellar and Uryasev (2013), highlighting the relationship between dispersion and risk measures. Moreover, in order to improve the out-of-sample performance of the portfolios when dealing with large problem dimension, we introduce ridge and elastic-net regularization. We propose quadratic programming formulations for the regularized problems, and we introduce a data-driven cross-validation procedure for the tuning of the parameter λ that regulates the strength of the penalty. Finally, we test the performance of the models on simulated and real-world data.

Simulations show that elastic-net and ridge penalties are effective in improving the out-of-sample performance compared to the unpenalized allocations, especially in presence of a large number of candidate assets. Also, the tuning procedure for the parameter λ is effective in selecting the optimal level of regularization. Finally, we observe that the differences in terms of performance among models with alternative values for the elastic-net parameter α are less relevant, with only a small advantage for the portfolios with $\alpha = 0.25$.

We finally test the penalized and non-penalized asset allocations in a rolling analysis framework, using a large dataset of US stocks. The allocations are compared to the traditional Markowitz portfolio and the equally weighted portfolio. The results are aligned to the simulation study, showing that the proposed minimum deviation allocations allow one to reduce out-of-sample portfolio risk and that they strongly benefit from the introduction of regularization on the weights. Robustness checks for different periods and datasets confirm the main results, suggesting that the regularization techniques are particularly useful for portfolios built using a large set of assets. Particularly interesting is the out-of-sample control of the left tail which leads to lower maximum drawdown of the regularized portfolios. This provides further support to previous studies, which however mostly focus on the minimum variance framework.

Our work opens new interesting lines of research. From a methodological point of view, we aim to further extend the tuning procedure of the elastic-net penalization, by simultaneously selecting both the parameters λ and α. Moreover, considering the portfolio replication framework, we plan to focus on creating sparse cloning strategies capable of better controlling asymmetric tail risk.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Gabriele Torri and Rosella Giacometti acknowledge the support of the Czech Science Foundation (GACR) under Project 19-11965S. Sandra Paterlini gratefully acknowledges financial support from DEM2019- QM-Lab-Emerging methods and models for high-dimensional and geo-coded data with applications in finance and economics.

ORCID

R. Giacometti http://orcid.org/0000-0002-9761-3529
G. Torri https://orcid.org/0000-0002-4934-2154
S. Paterlini https://orcid.org/0000-0003-4269-4496

References

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D., Coherent measures of risk. *Math. Finance*, 1999, **9**(3), 203–228.
Ban, G.-Y., El Karoui, N. and Lim, A.E., Machine learning and portfolio optimization. *Manage. Sci.*, 2016, **64**(3), 1136–1154.
Bassett, G.W., Koenker, R. and Kordas, G., Pessimistic portfolio allocation and choquet expected utility. *J. Financ. Econom.*, 2004, **2**(4), 477–492.
Bellini, F. and Bignozzi, V., On elicitable risk measures. *Quant. Finance*, 2015, **15**(5), 725–733.
Bellini, F. and Di Bernardino, E., Risk management with expectiles. *Eur. J. Finance*, 2017, **23**(6), 487–506.
Bellini, F., Mercuri, L. and Rroji, E., Implicit expectiles and measures of implied volatility. *Quant. Finance*, 2018, **18**(11), 1851–1864.
Belloni, A. and Chernozhukov, V., $\ell 1$-penalized quantile regression in high-dimensional sparse models. *Ann. Statist.*, 2011, **39**(1), 82–130.
Bonaccolto, G., Caporin, M. and Paterlini, S., Asset allocation strategies based on penalized quantile regression. *Comput. Manag. Sci.*, 2018, **15**(1), 1–32.
Broadie, M., Computing efficient frontiers using estimated parameters. *Ann. Oper. Res.*, 1993, **45**(1), 21–58.
Chopra, V.K. and Ziemba, W.T., The effect of errors in means, variances, and covariances on optimal portfolio choice. *J. Portf. Manag.*, 1993, **19**(2), 6–11.
Cont, R., Empirical properties of asset returns: Stylized facts and statistical issues. *Quant. Finance*, 2001, **1**(2), 223–236.
Danielsson, J. and Zhou, C., Why risk is so hard to measure. De Nederlandsche Bank Working Paper, No. 494, 2016.
DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Rev. Financ. Stud.*, 2007, **22**(5), 1915–1953.
Fan, J., Zhang, J. and Yu, K., Vast portfolio selection with gross-exposure constraints. *J. Am. Stat. Assoc.*, 2012, **107**(498), 592–606.
Fastrich, B., Paterlini, S. and Winker, P., Constructing optimal sparse portfolios using regularization methods. *Comput. Manag. Sci.*, 2015, **12**(3), 417–434.
Gendreau, B., Jin, Y., Nimalendran, M. and Zhong, X., CVaR-LASSO enhanced index replication (CLEIR): Outperforming by minimizing downside risk. *Appl. Econ.*, 2019, **51**(52), 5637–5651.
Giuzio, M. and Paterlini, S., Un-diversifying during crises: Is it a good idea? *Comput. Manag. Sci.*, 2019, **16**(3), 401–432.
Goto, S. and Xu, Y., Improving mean variance optimization through sparse hedging restrictions. *J. Financ. Quant. Anal.*, 2015, **50**(6), 1415–1441.
Gotoh, J. and Shinozaki, K., A robust portfolio technique for mitigating the fragility of CVAR minimization. Technical report, Dept. of Industrial and Systems Engineering, Chuo University, 2010.
Hu, W. and Zheng, Z., View bias towards ambiguity, expectile CAPM and the anomalies. Working paper, 2013.
Jagannathan, R. and Ma, T., Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance*, 2003, **58**(4), 1651–1683.

Koenker, R. and Bassett, Jr., G., Regression quantiles. *Econometrica*, 1978, **46**(1), 33–50.
Kremer, P.J., Lee, S., Bogdan, M. and Paterlini, S., Sparse portfolio selection via the sorted ℓ_1-norm. *J. Bank. Financ.*, 2020, **110**, 105687.
Laloux, L., Cizeau, P., Potters, M. and Bouchaud, J.-P., Random matrix theory and financial correlations. *Int. J. Theor. Appl. Finance*, 2000, **3**(3), 391–397.
Ledoit, O. and Wolf, M., A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.*, 2004, **88**(2), 365–411.
Ledoit, O. and Wolf, M., Robust performance hypothesis testing with the sharpe ratio. *J. Empir. Finance*, 2008, **15**(5), 850–859.
Ledoit, O. and Wolf, M., Robust performances hypothesis testing with the variance. *Wilmott*, 2011, **2011**(55), 86–89.
Liao, L., Park, C. and Choi, H., Penalized expectile regression: an alternative to penalized quantile regression. *Ann. Inst. Stat. Math.*, 2019, **71**(2), 409–438.
Markowitz, H., Portfolio selection. *J. Finance*, 1952, **7**(1), 77–91.
Michaud, R.O., The markowitz optimization enigma: Is 'optimized' optimal? *Finan. Anal. J.*, 1989, **45**(1), 31–42.
Newey, W.K. and Powell, J.L., Asymmetric least squares estimation and testing. *Econometrica*, 1987, **55**(4), 819–847.
Pflug, G.C., Pichler, A. and Wozabal, D., The 1/N investment strategy is optimal under high model ambiguity. *J. Bank. Financ.*, 2012, **36**(2), 410–417.
Rockafellar, R.T. and Uryasev, S., Optimization of conditional value-at-risk. *J. Risk*, 2000, **2**(3), 21–41.
Rockafellar, R.T. and Uryasev, S., The fundamental risk quadrangle in risk management, optimization and statistical estimation. *Surv. Oper. Res. Manag. Sci.*, 2013, **18**(1–2), 33–53.
Rockafellar, R.T., Uryasev, S. and Zabarankin, M., Risk tuning with generalized linear regression. *Math. Oper. Res.*, 2008, **33**(3), 712–729.
Salahi, M., Mehrdoust, F. and Piri, F., CVaR robust mean-CVaR portfolio optimization. *ISRN Appl. Math.*, 2013, **2013**, 1–9.
Sarykalin, S., Serraino, G. and Uryasev, S., Value-at-risk vs. conditional value-at-risk in risk management and optimization. In *State-of-the-Art Decision-Making Tools in the Information-Intensive Age*, pp. 270–294, 2008 (Informs). https://doi.org/10.1287/educ.1080.0052.
Tibshirani, R., Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B*, 1996, **58**(1), 267–288.
Torri, G., Giacometti, R. and Paterlini, S., Sparse precision matrices for minimum variance portfolios. *Comput. Manag. Sci.*, 2019, **16**(3), 375–400.

Appendix A. Quandrangles with expectiles

The *risk quadrangle* proposed by Rockafellar and Uryasev (2013) is a framework that relates the definition of risk measures, deviation measures and statistical estimation, in a scheme connected to generalized regression. In particular, the framework allows to define a set of four items (a risk measure ($\mathcal{D}(X)$), a deviation measure $\mathcal{R}(X)$, an error function $\mathcal{E}(X)$ and a regret function $\mathcal{V}(X)$), starting from a statistic $\mathcal{S}(X)$. We introduce here the quadrangle constructed using the expectile $EVaR_{1-\tau}$ as statistic. To our knowledge, this quadrangle has never been considered in the portfolio literature, even though the deviation measure proposed here corresponds to the *variancile* described in Hu and Zheng (2013).†

We underline that in the quadrangle proposed here the expectile is not used as a risk measures, but only as a statistic. The corresponding risk measure $\mathcal{R}$, as well as the other measures, are defined as follows:

- $\mathcal{R}_{\eta_\tau} = \mathbb{V}_\tau[X] - \mathbb{E}[X] = \min_{\xi\in\mathbb{R}} \mathbb{E}[\eta_\tau(X-\xi)] - \mathbb{E}[X]$,
- $\mathcal{D}_{\eta_\tau} = \mathbb{V}_\tau[X] = \min_{\xi\in\mathbb{R}} \mathbb{E}[\eta_\tau(X-\xi)]$,
- $\mathcal{E}_{\eta_\tau} = \mathbb{E}[\eta_\tau(X)]$
- $\mathcal{V}_{\eta_\tau} = \mathbb{E}[\eta_\tau(X)] - \mathbb{E}[X]$,

where $\eta_\tau(X) = X_+^2 + \frac{1-\tau}{\tau}X_-^2$, and the expectile $EVaR_{1-\tau}$ is defined implicitly as

$$EVaR_{1-\tau}(X) = -\arg\min_{\xi\in\mathbb{R}} \mathbb{E}[\eta_\tau(X-\xi)]. \tag{A1}$$

Note that for $\tau = 0.5$ the measure $\mathcal{D}_{\eta_\tau}$ is equal to the variance.

Appendix B. Minimum deviation portfolio as generalized regression

The global minimum deviation portfolio can be recasted as a generalized regression model with a loss function $\rho(X)$, in which the dependent variable is the nth asset, and there are $n-1$ regressors such that $R_i^* = R_n - R_i$:

$$\min_{\boldsymbol{w}_{-n}\in\mathbb{R}^{n-1},\xi\in\mathbb{R}} \mathbb{E}[\zeta_\tau(R_n - R^*w_1 - R^*w_2 - \cdots - R^*w_{n-1} - \xi)] \tag{A2}$$

Using the fact that $\sum_{i=1}^n w_i = 1$, we can then obtain the nth weight as $w_n = 1 - \sum_{i=1}^{n-1} w_i$.

Using the function $\zeta_\tau(X) = \rho_\tau(X) = X_+ + ((1-\tau)/\tau)X_-$, we have a quantile regression. Instead, if we use $\zeta_\tau(X) = \eta_\tau(X) = X_+^2 + ((1-\tau)/\tau)X_-^2$ we have expectile regression, while for $\zeta_\tau(X) = X^2$ we have the ordinary least square regression.

This result generalizes the ones for ordinary least square regression (Fan *et al.* 2012) and quantile regression (Bassett *et al.* 2004). On the practical level, it allows to use the very efficient algorithms developed for regression to solve portfolio optimization. Moreover, this approach is not suitable for penalized asset allocations (as it requires the choice of a numeraire, that is not penalized), and the introduction of additional constraints to the asset allocation is often extremely challenging, if not impossible.

Appendix C. Performance and portfolio measures

We describe the measures used to assess the portfolios performances, reported in tables 2–5 and A1–A4.

- $EVaR_{1-\tau}(X)$—Expectile risk measure with $\tau = 0.1$, computed as in equation (7). The indicator refers to a daily horizon.
- $CVaR_{1-\tau}(X)$—Conditional value at risk with $\tau = 0.1$, computed as in Rockafellar and Uryasev (2000). The indicator refers to a daily horizon.
- *Std*(X)—standard deviation of the portfolio returns, computed as $\mathrm{Std}(X) = \sqrt{\mathbb{E}[X^2 - \mathbb{E}[X]]}$. The indicator is annualized.
- $\sqrt{\mathcal{D}_{\eta_\tau}(X)}$—Square root of the deviation measure (6), with $\tau = 0.1$. The indicator is annualized.
- *Sharpe*(X)—Sharpe ratio of the portfolio, computed as $(\mu_X - r)/\sigma_X$, where μ_X is the return of the portfolio, r is the risk free rate (3 months treasury bill rate), and σ_X the standard deviation of the portfolio. The indicator is annualized.
- *max DD*—maximum drawdown. Computed as the maximum decline of a time series, from a peak to a nadir over a period of time.
- *Turnover*—average turnover of the portfolio. For each recalibration period, it is computed as $\sum_{j=1}^n |w_{j,t_k} - w_{j,t_{k-1}}|$, where n is the number of assets, t_k is the time of the portfolio rebalancing, t_{k-1} the previous one, and $w_{j,t}$ is the weight of the jth security at time t.

† A quadrangle built on the quantile as statistic $\mathcal{S}$, related to the quantile-based asset allocation describe above, can be found in Rockafellar and Uryasev (2013, section 2, example 2).

Table A1. Risk and performance measures for different asset allocation strategies. All the values are computed out--of-sample in a rolling windows scheme for the period 01/02/2010–01/02/2019. Datasets: SP100, FF48 τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. The results refer to the global minimum deviation portfolios (no constraints on the expected returns). The best and second best results for each measures are highlighted in bold and italics, respectively. The performance measures are described in Appendix 1.

	EVaR (%)	CVaR (%)	Std (%)	$\sqrt{\mathcal{D}_{\eta_\tau}}$ (%)	Sharpe	max DD (%)
Panel A—global minimum deviation portfolio (GMD)—SP100—2010–19						
MQ	0.665	1.45	12.8	21.9	0.504	15.2
MQ-R	0.552	1.25	10.7	19.3	0.855	13.1
MQ-EN	0.558	1.25	10.7	19.2	0.766	13.0
ME	0.583	1.3	11.3	19.8	0.804	10.9
ME-R	*0.545*	1.24	10.6	19.1	**0.91**	**12.2**
ME-EN	**0.545**	**1.23**	**10.5**	*19.0*	0.819	10.9
MV	0.555	1.24	10.7	**18.9**	0.816	*12.3*
MV-R	0.547	*1.24*	*10.6*	19.0	*0.856*	14.7
EW	0.862	1.96	16.1	29.4	0.641	24.6
Panel B—global minimum deviation portfolio (GMD)—FF48—2010–19						
MQ	0.564	1.27	11.4	19.6	**1.26**	14.5
MQ-R	0.535	1.22	10.6	18.7	1.17	14.7
MQ-EN	0.543	1.24	10.8	19.3	1.16	*13.8*
ME	0.541	1.22	10.9	18.8	1.19	14.1
ME-R	**0.524**	*1.19*	*10.4*	*18.4*	*1.21*	14.5
ME-EN	0.536	1.22	10.6	18.8	1.06	14.8
MV	*0.526*	**1.18**	**10.4**	**18.1**	1.13	**13.0**
MV-R	0.548	1.25	10.8	19.2	1.11	16.1
EW	0.845	1.93	16.2	29.0	0.79	24.0

Table A2. Portfolio statistics for different asset allocation strategies. All the values are computed out--of-sample in a rolling windows scheme for the period 01/02/2010–01/02/2019. Datasets: SP100, FF48 τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. The results refer to the global minimum deviation portfolios (no constraints on the expected returns). The performance measures are described in appendix 1.

	Turnover (%)	Gross (%)	Density (%)	max exp (%)
Panel A—global minimum deviation portfolio (GMD)—SP100—2010–19				
MQ	227.0	466.0	98.5	26.3
MQ-R	41.4	175.0	97.1	6.29
MQ-EN	47.2	134.0	45.4	11.9
ME	99.4	377.0	98.2	22.9
ME-R	35.1	180.0	97.7	6.01
ME-EN	31.8	145.0	84.5	10.4
MV	74.7	333.0	97.7	19.9
MV-R	24.0	165.0	97.2	5.52
Panel B—Global minimum deviation portfolio (GMD)—FF48—2010–19				
MQ	189.0	511.0	99.3	41.8
MQ-R	48.2	211.0	99.1	13.3
MQ-EN	55.5	171.0	56.6	20.2
ME	94.7	457.0	99.2	40.1
ME-R	33.6	211.0	98.8	13.3
ME-EN	25.4	157.0	85.3	18.4
MV	70.4	404.0	98.9	35.1
MV-R	14.5	165.0	98.8	9.68

- *Gross*—average gross exposure of the portfolio, computed as $\sum_{j=1}^{n} |w_j|$, where w_j is the weight of the jth security.
- *max exp.*—average maximum exposure of the portfolio on an individual asset, in absolute value. Computed as $\max_j |w_j|$, where w_j is the weight of the jth security.

Appendix D. Efficient frontiers

We compute the in-sample and out-of-sample efficient frontiers. The analysis is conducted using the same rolling window scheme proposed in section 5. In line with 2.3, for each asset allocation we compute both risk-return and the deviation-return frontiers. The deviation measures $\mathcal{D}_{\rho_\tau}(X)$ (quantile-based) and $\mathcal{D}_{\eta_\tau}(X)$ (expectile-based), and the risk measures associated in the risk-quadrangle framework ($\mathcal{R}_{\rho_\tau}(X) = CVaR_{1-\tau}(X)$, and $\mathcal{R}_{\eta_\tau}(X)$, respectively). For brevity, the analysis is conducted on the FF48 dataset, and considering the unpenalized asset allocations only.

Figure 1 shows the in-sample (solid blue) and out-of-sample (dashed orange) efficient frontiers for the FF48 dataset. The panels on the left show the deviation-return frontier, the ones in the center the risk-return frontier, and the ones on the right the variance-return frontier. The in-sample and out-of-sample global minimum deviation portfolios and global minimum risk portfolios are highlighted on the plot as asterisks and circles, respectively.

We see that the shape of the out-of-sample frontiers is quite different from the in-sample ones. In particular, the expected returns of the portfolio on the frontier are much smaller, consistently with a bias in the estimation of the in-sample efficient frontier (Broadie 1993).

Considering the deviation-return frontiers (left plots), we see that the out-of-sample global minimum deviation portfolio is, as expected, on the left part of the frontier, while the minimum risk has a higher deviation, and a higher return. The risk-return frontiers (right plots) shows instead that out-of-sample the minimum deviation portfolios are characterized by a smaller risk compared to the global minimum risk portfolios. This is not consistent with the in-sample results and is probably due to the larger impact of the estimation of expected returns for the minimum risk portfolios compared to the minimum deviation ones.

Overall, results are consistent with previous literature on portfolio optimization, suggesting that the portfolio weights are particularly sensitive to the estimation of returns, not only in the mean–variance framework but also in the quantile and expectile asset allocations proposed here.

Appendix E. Robustness checks

We test the robustness of the results by considering two additional datasets: the set of 91 stocks included in the Standard and Poor's 100 equity index (SP100), and the Fama and French 48 industry portfolio (FF48) for the investment period 01/02/2010–01/02/2019. The results for the GMD portfolios are reported in table A1.† We see that overall the results are aligned to the ones reported for the SP500 dataset. The main difference is that in these set-ups, and especially for the FF48 portfolios, the performances of the unpenalized approach are more aligned to the ones of the penalized portfolios, most likely due to a smaller effect of estimation error when the number of assets is smaller. The results underlay that such techniques are particularly suitable for dealing with portfolios constructed with a

† The results for MD and GMD-LO are omitted for brevity.

Table A3. Risk and performance measures for different asset allocation strategies. All the values are computed out-of-sample in a rolling windows scheme for the period 01/02/2012–01/02/2019 (1000 daily observation in each estimation window). Datasets: SP500, SP100, FF48. τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. The results refer to the global minimum deviation portfolios (no constraints on the expected returns). The best and second best results for each measures are highlighted in bold and italics, respectively. The performance measures are described in Appendix 1.

	EVaR (%)	CVaR (%)	Std (%)	$\sqrt{\mathcal{D}_{\eta_\tau}}$ (%)	Sharpe	max DD (%)
Panel A—global minimum deviation portfolio (GMD)—SP500—2012–19—1000 obs.						
MQ	0.648	1.39	12.3	20.6	0.34	27.6
MQ-R	*0.438*	*0.99*	8.86	*15.5*	*1.07*	11.6
MQ-EN	0.467	1.05	9.27	16.3	0.96	12.6
ME	0.539	1.16	10.4	17.4	0.55	12.8
ME-R	**0.436**	**0.98**	**8.66**	**15.2**	1.01	*11.5*
ME-EN	0.452	1.02	8.91	15.9	0.99	**11.4**
MV	0.514	1.12	9.99	16.7	0.69	14.4
MV-R	0.441	1.00	*8.79*	15.7	**1.12**	12.2
Panel B—global minimum deviation portfolio (GMD)—SP100—2012–19—1000 obs.						
MQ	0.574	1.26	11.2	19.1	0.72	24.6
MQ-R	*0.533*	1.20	*10.4*	18.4	*0.92*	14.3
MQ-EN	0.542	1.21	10.5	18.6	0.83	16.4
ME	0.560	1.23	10.8	18.5	0.57	24.6
ME-R	**0.523**	**1.18**	**10.3**	*18.2*	**1.01**	*14.3*
ME-EN	0.543	1.22	10.4	18.5	0.80	**13.7**
MV	0.539	*1.19*	10.4	**18.0**	0.62	16.8
MV-R	0.557	1.26	10.6	19.0	0.82	15.0
Panel C—global minimum deviation portfolio (GMD)—FF48—2012-19—1000 obs.						
MQ	0.540	1.21	10.8	18.5	1.07	14.0
MQ-R	0.522	1.19	10.2	18.0	1.18	14.2
MQ-EN	0.542	1.23	10.5	18.5	0.95	14.3
ME	0.525	*1.18*	10.5	18.0	1.14	*13.0*
ME-R	*0.519*	1.19	**10.2**	*18.0*	*1.19*	13.5
ME-EN	0.532	1.20	10.4	18.2	1.07	12.9
MV	**0.505**	**1.14**	*10.2*	**17.5**	**1.27**	**12.3**
MV-R	0.527	1.20	10.4	18.3	1.19	14.9
EW	0.716	1.62	13.6	23.8	0.85	21.1

very large number of assets, where the estimation error plays a more prominent role. Since the assets in the portfolios SP100 are a subset of the ones for SP500, we can compare directly the two portfolios. We see that the penalized models take advantage of the increased number of assets and obtain smaller risk, in terms of all the measures considered, for both quantile and expectile allocations. On the contrary, the performances of the unpenalized portfolios are significantly worse for SP500 compared to SP100, as the estimation error has a stronger impact due to the larger number of assets. Concerning portfolio statistics, reported in table A2, the results confirm the ones reported in the text, showing how the introduction of penalization allows to reduce turnover and gross exposures.

Table A4. Portfolio statistics for different asset allocation strategies. All the values are computed out-of-sample in a rolling windows scheme for the period 01/02/2012–01/02/2019 (1000 daily observation in each estimation window). Datasets: SP500, SP100, FF48. τ is always equal to 0.1. For the elastic-net penalty, the parameter α has been set to 0.25. The results refer to the global minimum deviation portfolios (no constraints on the expected returns). The portfolio measures are described in appendix 1.

	Turnover (%)	Gross (%)	Density (%)	max exp (%)
Panel A—global minimum deviation portfolio (GMD)—SP500—2012–19—1000 obs.				
MQ	649.0	1260.0	97.2	17.4
MQ-R	99.3	383.0	93.4	3.5
MQ-EN	115.0	254.0	36.2	7.58
ME	262.0	1010.0	96.6	13.6
ME-R	90.2	394.0	92.4	3.28
ME-EN	37.4	221.0	74.8	5.49
MV	218.0	942.0	96.3	13.2
MV-R	50.2	328.0	91.4	2.43
Panel B—global minimum deviation portfolio (GMD)—SP100—2012–19—1000 obs.				
MQ	114.0	372.0	98.0	21.2
MQ-R	39.6	168.0	97.1	6.26
MQ-EN	47.1	140.0	53.1	11.3
ME	50.8	336.0	97.5	19.6
ME-R	23.3	174.0	96.8	6.41
ME-EN	24.4	140.0	86.7	9.54
MV	37.6	298.0	97.6	18.5
MV-R	14.0	144.0	97.3	4.63
Panel C—global minimum deviation portfolio (GMD)—FF48—2012–19—1000 obs.				
MQ	113.0	475.0	99.5	36.3
MQ-R	37.4	195.0	98.7	11.5
MQ-EN	44.5	159.0	59.8	16.0
ME	49.6	432.0	98.8	34.9
ME-R	24.2	194.0	98.6	11.6
ME-EN	19.7	142.0	81.0	14.6
MV	37.2	378.0	98.6	31.2
MV-R	9.9	160.0	99.0	8.9

As an additional robustness check, we test the model considering a longer estimation period, equal to 1000 daily observations for each window (approximately 4 years), in order to test the sensitivity to the size of the estimation sample. Table A3 reports the results for the SP500, SP100 and FF48 datasets. For brevity we report only the global minimum deviation (GMD) portfolios. The results overall are aligned to the data presented in the main text. More in detail, we see that the ridge penalized expectile based portfolios (ME-R) outperform the other in terms of most of the risk measures for the SP500 and SP100 dataset, while the minimum variance portfolio is the preferable one when considering the FF48 dataset. The portfolio statistics, reported in table A4 are also aligned with the results in the main text.

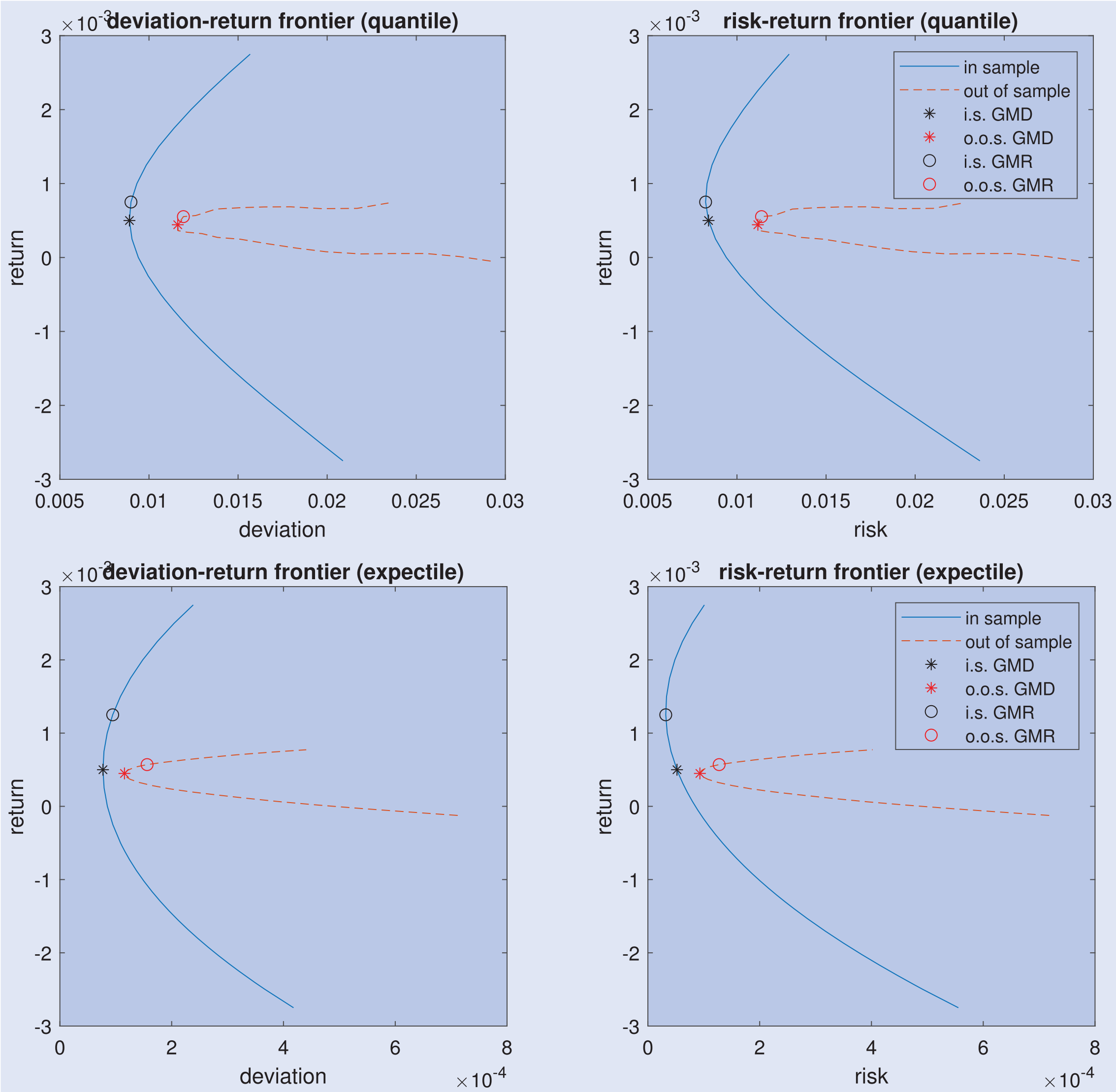

Figure A1. Efficient frontiers for expectile and quantile asset allocations. Left panels represent the deviation-return frontier, center panels risk-return frontier and right panels variance-return. Note that the risk and deviation measures are the ones corresponding to the respective optimization model (see the quadrangles plot above). The global minimum deviation (GMD) and global minimum risk (GMR) portfolios are highlighted as astericks and circles, respectively. The reported data are the average across all the estimation windows.