Applied Mathematics and Computation 427 (2022) 127120

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Bayesian portfolio selection using VaR and CVaR

Taras Bodnar*, Mathias Lindholm, Vilhelm Niklasson, Erik Thorsén

Department of Mathematics, Stockholm University, Stockholm SE-10691, Sweden

ARTICLE INFO

Article history:
Received 22 October 2021
Revised 26 January 2022
Accepted 22 March 2022
Available online 19 April 2022

Keywords:
Bayesian inference
Posterior predictive distribution
Optimal portfolio
VaR
CVaR

ABSTRACT

We study the optimal portfolio allocation problem from a Bayesian perspective using value at risk (VaR) and conditional value at risk (CVaR) as risk measures. By applying the posterior predictive distribution for the future portfolio return, we derive relevant quantities needed in the computations of VaR and CVaR, and express the optimal portfolio weights in terms of observed data only. This is in contrast to the conventional method where the optimal solution is based on unobserved quantities which are estimated. We also obtain the expressions for the weights of the global minimum VaR (GMVaR) and global minimum CVaR (GMCVaR) portfolios, and specify conditions for their existence. It is shown that these portfolios may not exist if the level used for the VaR or CVaR computation are too low. By using simulation and real market data, we compare the new Bayesian approach to the conventional plug-in method by studying the accuracy of the GMVaR portfolio and by analysing the estimated efficient frontiers. It is concluded that the Bayesian approach outperforms the conventional one, in particular at predicting the out-of-sample VaR.

1. Introduction

The economic theory underlying optimal portfolio selection was pioneered by Markowitz [28]. In his seminal paper, the optimal allocation of the available assets was determined by their expected returns and the covariance matrix of the asset returns. In practise, however, both the vector of expected returns and the covariance matrix are unknown and have to be estimated using historical data. These estimates were traditionally treated as the true parameters of the data-generating process and plugged into the equations for the weights of optimal portfolios. Unfortunately, this causes a challenging problem since the optimal portfolio weights appear to be sensitive to misspecification of the input parameters, especially the expected returns, and thus estimation errors can lead to poorly performing portfolios (see, Huang and Di [23], Merton [30], Muhinyuza et al. [31], Simaan [34]).

The problem concerning parameter uncertainty was the main reason why the Bayesian approach was introduced in portfolio theory during the 1970s. In the Bayesian setting, the parameters of the distribution of the asset returns are modelled as random variables with a prior distribution summarizing preexisting information. The posterior distribution, derived from the likelihood function and the used priors, provides updated knowledge on the model parameters conditioned on the observed data. Moreover, the posterior predictive distribution expresses the distribution of the next returns conditioned on the previous returns. This distribution has been widely used in Bayesian portfolio theory to account for parameter uncertainty (see,

* Corresponding author.
E-mail address: taras.bodnar@math.su.se (T. Bodnar).

https://doi.org/10.1016/j.amc.2022.127120

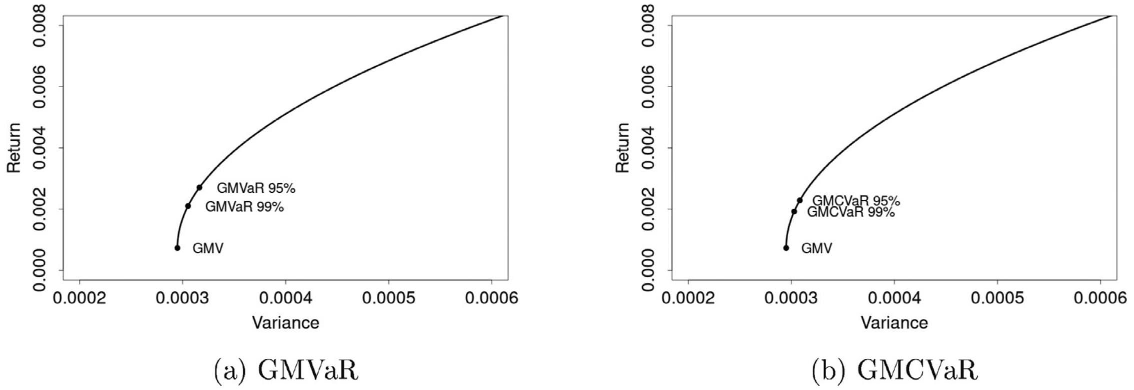


Fig. 1. Mean-variance efficient frontier based on empirical data from stocks in the S&P 500 index together with the locations of the GMV, GMVaR and GMCVaR portfolios.

e.g., Barry [5], Klein and Bawa [25], Winkler [38]). Recently, Bauder et al. [8] derived an expression of it in terms of two independent Student's t-distributions under the assumption of conditional normality and from that obtained the optimal mean-variance portfolio weights.

During the last decades it has become increasingly popular to use value at risk (VaR) and conditional value at risk (CVaR) instead of the variance in the portfolio optimization problem (see, e.g., Ortiz-Gracia and Oosterlee [32]). In this setting, the goal is to minimize the risk in terms of the portfolio VaR or CVaR for a certain level of expected return. VaR and CVaR are quantile-based risk measures focusing on downside risk, meaning that the risk is determined from a quantile in the right tail of a loss distribution. Alexander and Baptista [1,2] derived the weights of a portfolio when using VaR or CVaR in the objective function in the portfolio optimization problem under the assumption of multivariate normally distributed asset returns. They noted that the optimal allocation strategy does not depend on whether the variance, VaR or CVaR is used as a risk measure. However, the portfolio that globally minimizes VaR or CVaR may not coincide with the one which globally minimizes the variance. This is illustrated in Fig. 1 where the mean-variance efficient frontier is plotted together with the locations of the global minimum variance (GMV), global minimum VaR (GMVaR) and global minimum CVaR (GMCVaR) portfolios based on weekly returns for 20 randomly selected stocks in the S&P 500 index. In this figure, the efficient frontier and the optimal portfolios are all calculated using the conventional method, i.e., using sample estimates and treating them as the true parameters of the data-generating model. The fact that the GMVaR and GMCVaR portfolios are located above the GMV portfolio on the mean-variance efficient frontier motivates the study of these portfolios also for an investor who constructs the portfolio following the mean-variance analysis, since their expected returns determine the lower bounds for an investor who wants to be efficient also from a mean-VaR or mean-CVaR perspective. Given the Basel regulations that require banks to use VaR and CVaR in their risk reporting, being efficient also from these perspectives should be highly desirable.

This paper contributes to the current literature on portfolio theory by combining a Bayesian framework and an asset allocation strategy based on VaR and CVaR. On the one hand, we extend the results of Alexander and Baptista [1,2] by taking the parameter uncertainty into account, while, on the other hand, we develop the results of Bauder et al. [8], who derived the expressions of the mean-variance optimal portfolio weights from a Bayesian perspective. Similarly to Bauder et al. [8], the posterior predictive distribution is used in the derivation of the optimization problem. However, the portfolio variance is replaced by VaR and CVaR in the derivation. We further compare the new Bayesian approach to its frequentist counterpart and demonstrate some of the advantages of the former through a simulation study and an empirical illustration.

The paper is organized as follows. In Section 2, we present a Bayesian model of portfolio returns and derive a stochastic representation of its posterior predictive distribution. In Section 3, we use the obtained stochastic representation to establish expressions for VaR and CVaR of a portfolio. Section 4 deals with portfolio optimization under parameter uncertainty. The Bayesian portfolio optimization problem is solved using VaR and CVaR as risk measures and conditions for the existence of a GMVaR and GMCVaR portfolio are presented. Moreover, the mean-VaR and mean-CVaR efficient frontiers are derived. Section 5 contains a simulation study where our Bayesian approach is compared to the conventional method by analysing the accuracy of the GMVaR portfolio as well as by comparing the estimated efficient frontiers. The comparison is continued in Section 6 with an empirical study based on real market data. Section 7 contains a conclusion and discussion based on our findings. All proofs can be found in the appendix together with a description of the data.

2. A Bayesian model of portfolio and asset returns

In this section, we introduce our Bayesian model and derive the posterior predictive distributions of the portfolio return using two different priors.

2.1. Posterior predictive distribution of portfolio return

Let $\mathbf{X}_t$ be a k-dimensional vector of logarithmic asset returns at time t, i.e., each element $X_{t,i}$, $i = 1, \ldots, k$ of $\mathbf{X}_t$ is defined as

$$X_{t,i} := \log\left(\frac{P_{t,i}}{P_{t-1,i}}\right),$$

where $P_{t,i}$ denotes the price at time t of asset i. Moreover, let $\mathbf{x}_{(t-1)} = (\mathbf{x}_{t-n}, \ldots, \mathbf{x}_{t-1})$ be the observation matrix of the asset returns $\mathbf{x}_{t-n}$, ..., $\mathbf{x}_{t-1}$ taken from time $t-n$ until $t-1$ whose distribution depends on the parameter vector $\boldsymbol{\theta}$. Bayes' theorem provides the posterior distribution of $\boldsymbol{\theta}$ expressed as

$$\pi(\boldsymbol{\theta}|\mathbf{x}_{(t-1)}) \propto f(\mathbf{x}_{(t-1)}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}),$$

where $\pi(\cdot)$ is the prior distribution and $f(\cdot|\boldsymbol{\theta})$ is the likelihood function.

At time t the return of the portfolio with weights $\mathbf{w} = (w_1, \ldots, w_k)$, where it is assumed that $\mathbf{1}^\top \mathbf{w} = 1$, is given by

$$X_{P,t} := \mathbf{w}^\top \mathbf{X}_t.$$

The posterior predictive distribution of $X_{P,t}$, i.e., the conditional distribution of $X_{P,t}$ given $\mathbf{x}_{(t-1)}$, is computed by (see, e.g., p. 244 in Bernardo and Smith [9])

$$f(x_{P,t}|\mathbf{x}_{(t-1)}) = \int_{\boldsymbol{\theta}\in\Theta} f(x_{P,t}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathbf{x}_{(t-1)})d\boldsymbol{\theta}, \tag{2.1}$$

where Θ denotes the parameter space and $f(\cdot|\boldsymbol{\theta})$ is the conditional density of $X_{P,t}$ given $\boldsymbol{\theta}$. The posterior predictive distribution (2.1) determines the distribution of the future portfolio return at time t given the information available in the historical data of asset returns up to $t-1$. It can be used to construct a point prediction of the future portfolio return, such as the posterior predictive mean or mode, together with the uncertainty, expressed as the posterior predictive variance. Furthermore, a prediction interval for the portfolio return can be obtained as a posterior predictive credible interval.

Employing the non-informative Jeffreys prior and the informative conjugate prior, Bauder et al. [8] derived stochastic representations of a random variable following the posterior predictive distribution (2.1) under the assumption that $\mathbf{X}_1, \mathbf{X}_2, \ldots$ are conditionally i.i.d. and $\mathbf{X}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The Jeffreys prior and the conjugate prior are then given by

$$\pi_J(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(k+1)/2}, \tag{2.2}$$

and

$$\boldsymbol{\mu}|\boldsymbol{\Sigma} \sim N_k\left(\mathbf{m}_0, \frac{1}{r_0}\boldsymbol{\Sigma}\right), \quad \text{and} \quad \boldsymbol{\Sigma} \sim IW_k(d_0, \mathbf{S}_0), \tag{2.3}$$

respectively, where $|\cdot|$ stands for the determinant and $IW_k(d_0, \mathbf{S}_0)$ denotes the inverse Wishart distribution with d_0 degrees of freedom and parameter matrix $\mathbf{S}_0$ (see, e.g., Section 3.4 in Gupta and Nagar [21], for the definition and properties). The quantities $\mathbf{m}_0$, r_0, d_0, and $\mathbf{S}_0$ are the hyperparameters of the conjugate prior where $\mathbf{m}_0$ and $\mathbf{S}_0$ reflect the investor's prior belief about the mean vector and covariance matrix, whereas r_0 and d_0 represent the precision of these beliefs. The two priors (2.2) and (2.3) have been widely used in financial literature where the Jeffreys prior is usually also referred to as the diffuse prior (see, e.g., Barry [5], Bodnar et al. [13], Brown [17], Stambaugh [35]), and the conjugate prior is also commonly related to the Black-Litterman model (see, Bauder et al. [8], Black and Litterman [10], Frost and Savarino [20], Kolm and Ritter [26]).

The stochastic representations derived by Bauder et al. [8] fully determine the posterior predictive distribution. They are expressed in terms of two independent standard t-distributed random variables whose degrees of freedom depend on the assigned prior. Using these findings, the posterior predictive distribution is deduced and is presented in Proposition 2.1. Let $t(q, a, b^2)$ denote the univariate t-distribution with q degrees of freedom, location parameter a and scale parameter b. Then, we obtain the following results.

Proposition 2.1. *Let asset returns $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be conditionally i.i.d. with $\mathbf{X}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then*

1. *under the Jeffreys prior, it holds for $n > k$ that the posterior predictive distribution defined in (2.1) is $t(d_{k,n,J}, \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,J}, r_{k,n,J}\mathbf{w}^\top \mathbf{S}_{t-1,J}\mathbf{w})$ with $d_{k,n,J} = n - k$,*

$$r_{k,n,J} = \frac{n+1}{n(n-k)}, \quad \bar{\mathbf{x}}_{t-1,J} = \frac{1}{n}\sum_{i=t-n}^{t-1} \mathbf{x}_i, \tag{2.4}$$

and

$$\mathbf{S}_{t-1,J} = \sum_{i=t-n}^{t-1} (\mathbf{x}_i - \bar{\mathbf{x}}_{t-1,J})(\mathbf{x}_i - \bar{\mathbf{x}}_{t-1,J})^\top; \tag{2.5}$$

2. *under the conjugate prior, it holds for* $n+d_0-2k>0$ *that the posterior predictive distribution defined in (2.1) is* $t(d_{k,n,C}, \mathbf{w}^\top\bar{\mathbf{x}}_{t-1,C}, r_{k,n,C}\mathbf{w}^\top\mathbf{S}_{t-1,C}\mathbf{w})$ *with* $d_{k,n,C}=n+d_0-2k$,

$$r_{k,n,C}=\frac{n+r_0+1}{(n+r_0)(n+d_0-2k)}, \quad \bar{\mathbf{x}}_{t-1,C}=\frac{n\bar{\mathbf{x}}_{t-1,J}+r_0\mathbf{m_0}}{n+r_0}, \tag{2.6}$$

and

$$\mathbf{S}_{t-1,C}=\mathbf{S}_{t-1,J}+\mathbf{S}_0+nr_0\frac{(\mathbf{m_0}-\bar{\mathbf{x}}_{t-1,C})(\mathbf{m_0}-\bar{\mathbf{x}}_{t-1,C})^\top}{n+r_0}. \tag{2.7}$$

The proof of Proposition 2.1 is given in the appendix. It is a proof based on the results of Bauder et al. [8], although other authors have provided similar results (see, e.g., Klein and Bawa [25], Winkler [38]).

Since the structure of the posterior predictive distributions are similar, we introduce new notation which allows us to combine the two findings of Proposition 2.1 into a single result. Let

$$\begin{cases} d_{k,n}=d_{k,n,J},\ r_{k,n}=r_{k,n,J},\ \bar{\mathbf{x}}_{t-1}=\bar{\mathbf{x}}_{t-1,J},\ \mathbf{S}_{t-1}=\mathbf{S}_{t-1,J}, & \text{under the Jeffreys prior,}\\ d_{k,n}=d_{k,n,C},\ r_{k,n}=r_{k,n,C},\ \bar{\mathbf{x}}_{t-1}=\bar{\mathbf{x}}_{t-1,C},\ \mathbf{S}_{t-1}=\mathbf{S}_{t-1,C}, & \text{under the conjugate prior.}\end{cases} \tag{2.8}$$

From Proposition 2.1 we deduce that the posterior predictive distribution under both priors can be expressed as $t(d_{k,n}, \mathbf{w}^\top\bar{\mathbf{x}}_{t-1}, r_{k,n}\mathbf{w}^\top\mathbf{S}_{t-1}\mathbf{w})$ with $d_{k,n}$, $r_{k,n}$, $\bar{\mathbf{x}}_{t-1}$, and $\mathbf{S}_{t-1}$ as in (2.8). Let $\widehat{X}_{P,t}$ denote a random variable which follows the posterior predictive distribution, i.e., whose distribution coincides with the conditional distribution of portfolio return $X_{P,t}$ given the information available up to time $t-1$. Application of Proposition 2.1 together with (2.8) then gives the following stochastic representation of $\widehat{X}_{P,t}$

$$\widehat{X}_{P,t}\overset{d}{=}\mathbf{w}^\top\bar{\mathbf{x}}_{t-1}+\tau\sqrt{r_{k,n}}\sqrt{\mathbf{w}^\top\mathbf{S}_{t-1}\mathbf{w}}, \tag{2.9}$$

where $\tau\sim t(d_{k,n})$ with $t(d_{k,n})$ denoting the standard t-distribution with $d_{k,n}$ degrees of freedom, i.e., the t-distribution with zero location parameter and scale parameter equal to one. Representation (2.9) immediately implies that

$$\mathrm{E}[\widehat{X}_{P,t}]=\mathrm{E}[X_{P,t}|\mathbf{x}_{(t-1)}]=\mathbf{w}^\top\bar{\mathbf{x}}_{t-1},\ \mathrm{Var}(\widehat{X}_{P,t})=\mathrm{Var}(X_{P,t}|\mathbf{x}_{(t-1)})=\frac{d_{k,n}r_{k,n}}{d_{k,n}-2}\mathbf{w}^\top\mathbf{S}_{t-1}\mathbf{w}, \tag{2.10}$$

for $d_{k,n}>1$ and $d_{k,n}>2$, respectively.

2.2. Time series representation of asset returns

In the financial literature it is common to describe asset returns in terms of a time series model (see, e.g., Tsay [36]). In order to make comparisons to such models easier, we now give a time series representation of the considered Bayesian model:

$$\mathbf{X}_1,\mathbf{X}_2,\dots,\mathbf{X}_t|\boldsymbol{\mu},\boldsymbol{\Sigma}\ \text{ are independent with }\ \mathbf{X}_i|\boldsymbol{\mu},\boldsymbol{\Sigma}\sim N_k(\boldsymbol{\mu},\boldsymbol{\Sigma}). \tag{2.11}$$

Model (2.11) should not be confused with the assumption that the asset returns are independent and normally distributed as it is usually presented in financial and econometric literature. The two models coincide only when $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ possess probability distributions concentrated in single points. This would imply that the investor knows $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with certainty, which is not the case in practice.

In order to show the considerable difference between the model (2.11) and the model assuming independent and normally distributed asset returns, we next derive the marginal distribution of $\mathbf{X}_1$, $\mathbf{X}_2$, ..., $\mathbf{X}_t$ by using the findings of Proposition 2.1 and the stochastic representation (2.9). The results will be obtained by assigning the Jeffreys prior to the model parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ as well as the conjugate prior.

In Proposition 2.1 we showed that given a sample of size $n>k$, the posterior predictive distribution of $X_{P,t}=\mathbf{w}^\top\mathbf{X}_t$ given $\mathbf{X}_{t-n},\dots,\mathbf{X}_{t-1}$ has a univariate t-distribution under both priors (see, also the stochastic representation (2.9)). Since $\mathbf{w}$ is an arbitrary vector with the only restriction $\mathbf{1}^\top\mathbf{w}=1$, choosing $n=t-1$ we get

$$\mathbf{X}_t|\mathbf{X}_1,\dots,\mathbf{X}_{t-1}\sim t_k(d_{k,t-1},\bar{\mathbf{x}}_{1:t-1},r_{k,t-1}\mathbf{S}_{1:t-1}), \tag{2.12}$$

and, similarly for $j=k+2,\dots,t-1$ it holds that

$$\mathbf{X}_{j+1}|\mathbf{X}_1,\dots,\mathbf{X}_j\sim t_k(d_{k,j},\bar{\mathbf{x}}_{1:j},r_{k,j}\mathbf{S}_{1:j}), \tag{2.13}$$

where $d_{k,j}$ and $r_{k,j}$ for $j=k+2,\dots,t-1$ are defined as in (2.8); $\bar{\mathbf{x}}_{1:j}$ and $\mathbf{S}_{1:j}$ are defined similarly to $\bar{\mathbf{x}}_{t-1}$ and $\mathbf{S}_{t-1}$, where the sums in (2.4) and (2.5) are from 1 to j. The symbol $t_k(\cdot,\cdot,\cdot)$ stands for the k-dimensional multivariate t-distribution. The last result can also be written in the time series context in the following way

$$\mathbf{X}_{j+1}=\bar{\mathbf{x}}_{1:j}+\sqrt{r_{k,j}}\mathbf{S}_{1:j}^{1/2}\boldsymbol{\varepsilon}_j \quad \text{for} \quad j=k+2,\dots,t-1, \tag{2.14}$$

where $\boldsymbol{\varepsilon}_j\sim t_k(d_{k,j},\mathbf{0},\mathbf{I})$ is the error term and $\mathbf{S}_{1:j}^{1/2}$ denotes a square root of $\mathbf{S}_{1:j}$.

The derived time series model for $\mathbf{X}_1, \ldots, \mathbf{X}_t$ indicates that a complicated nonlinear dependence structure is present in the conditional model (2.11). Moreover, model (2.11) can be used to capture the time-dependent structure in the dynamics of both the conditional mean vector and the conditional covariance matrix of the asset returns which is usually observed in practice (see, e.g., Tsay [36]). Finally, the error terms in the time series model for $\mathbf{X}_1, \ldots, \mathbf{X}_t$ in (2.14) are multivariate t-distributed with degrees of freedom depending on the amount of available information. This gives rise to an additional source of non-stationarity and allows for heavier tails than in a model where asset returns are assumed to be independent and normally distributed corresponding to the conventional approach.

3. Portfolio VaR and CVaR

The quantile based risk-measures value at risk (VaR) and conditional value at risk (CVaR) are introduced in the following section and we derive their analytical expressions in our Bayesian setting using the posterior predictive distribution.

3.1. Posterior predictive VaR and CVaR

The results of Proposition 2.1 provides an easy way to find the quantiles of the posterior predictive distribution. The two most common quantile-based risk measures used in the literature are VaR and CVaR. Denoting by $\widehat{X}_{P,t}$ a random variable whose distribution coincides with the posterior predictive distribution of the portfolio return $X_{P,t}$ given $\mathbf{x}_{(t-1)}$ and using that the posterior predictive distribution is absolutely continuous, the two quantile-based risk measures at level $\alpha \in (0.5, 1)$ are defined by

$$\mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t}) := F^{-1}_{Y,t-1}(\alpha)$$

and

$$\mathrm{CVaR}_{\alpha,t-1}(\widehat{X}_{P,t}) := \mathrm{E}[Y|Y \geq \mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t})],$$

respectively, where $Y := -\widehat{X}_{P,t}$ is the portfolio loss with cumulative distribution function $F_{Y,t-1}(\cdot)$.

In the following, it is implicit that the probability and expectation in the definitions of the VaR and CVaR are formulated in terms of the posterior predictive distribution and they are conditioned on previous asset returns $\mathbf{x}_{(t-1)}$. Note also that the posterior predictive distribution is free of parameters and is fully determined by the observed data.

By definition, $\mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t})$ satisfies

$$P\Big(\widehat{X}_{P,t} \leq -\mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t})\Big) = 1-\alpha. \tag{3.1}$$

Rewriting the left hand side of (3.1) yields

$$P\left(\mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \tau\sqrt{r_{k,n}}\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1}\mathbf{w}} \leq -\mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t})\right) = P\left(\tau \leq \frac{-\mathrm{VaR}_{\alpha}(\widehat{X}_{P,t}) - \mathbf{w}^\top \bar{\mathbf{x}}_{t-1}}{\sqrt{r_{k,n}}\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1}\mathbf{w}}}\right),$$

where τ is standard t-distributed with degrees of freedom $d_{k,n}$ defined in (2.8) depending on the prior. Hence,

$$\frac{-\mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t}) - \mathbf{w}^\top \bar{\mathbf{x}}_{t-1}}{\sqrt{r_{k,n}}\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1}\mathbf{w}}} = d_{1-\alpha},$$

where $d_{1-\alpha}$ is the $(1-\alpha)$ quantile of the t-distribution with $d_{k,n}$ degrees of freedom which satisfies $d_\alpha = -d_{1-\alpha}$ due to the symmetry of the t-distribution. Thus,

$$\mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t}) = -\mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + d_\alpha\sqrt{r_{k,n}}\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1}\mathbf{w}}. \tag{3.2}$$

Similarly, it follows by the definition of CVaR that

$$\mathrm{CVaR}_{\alpha,t-1}(\widehat{X}_{P,t}) = \mathrm{E}\big[-\widehat{X}_{P,t}| -\widehat{X}_{P,t} \geq \mathrm{VaR}_{\alpha,t-1}(\widehat{X}_{P,t})\big] = -\mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + k_\alpha\sqrt{r_{k,n}}\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1}\mathbf{w}}, \tag{3.3}$$

with

$$k_\alpha = \mathrm{E}[-\tau| -\tau \geq d_\alpha] = \frac{1}{1-\alpha}\int_{d_\alpha}^{\infty} t f_{d_{k,n}}(t)\mathrm{d}t = \frac{1}{1-\alpha}\frac{\Gamma\left(\frac{d_{k,n}+1}{2}\right)}{\Gamma\left(\frac{d_{k,n}}{2}\right)\sqrt{\pi d_{k,n}}}\frac{d_{k,n}}{d_{k,n}-1}\left(1+\frac{d_\alpha^2}{d_{k,n}}\right)^{-\frac{d_{k,n}-1}{2}},$$

where $f_{d_{k,n}}(t)$ denotes the density of the t-distribution with $d_{k,n}$ degrees of freedom and we use that the distribution of $-\tau$ coincides with τ due to the symmetry of the t-distribution.

The expressions of VaR and CVaR given in (3.2) and (3.3), respectively, can be presented in the following common form

$$Q_{t-1}(\mathbf{w}) = -\mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + q_\alpha \sqrt{r_{k,n}} \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}}, \tag{3.4}$$

where $Q_{t-1}(\mathbf{w})$ denotes either VaR or CVaR and q_α is equal to d_α when the former is considered and k_α when latter is considered. This general formulation will be used extensively in the rest of the paper in order to handle VaR and CVaR cases simultaneously.

Since $\alpha \in (0.5, 1)$ and the t-distribution is symmetric, we get that $d_\alpha > 0$. Moreover, we have that $k_\alpha > 0$ by definition. As a result, $q_\alpha > 0$ which together with the convexity of $\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}}$ implies the following result.

Proposition 3.1. *Under the conditions of Proposition 2.1, $Q_{t-1}(\mathbf{w})$ is convex with respect to* $\mathbf{w}$.

The proof of Proposition 3.1 is given in the appendix.

Remark 3.2. The stochastic representation in (2.9) shows that the general presentation of the posterior predictive VaR and CVaR as given in (3.4) can be extended to other risk measures used in portfolio theory and risk management. Such a result will provide a possibility to formulate and to solve the portfolio choice problem under more general setups considered in financial mathematics which is based on the concept of coherent risk measures (see, e.g., Artzner et al. [3]).

Assuming that the asset returns can be modelled as $\mathbf{X}_t | \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the analytical expression of the risk using a general coherent risk measure $\rho(X_{P,t})$ is deduced from (2.9) and it is given by

$$\rho_{t-1}(\widehat{X}_{P,t}) = \rho_{t-1}(\mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \tau \sqrt{r_{k,n}} \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}}) = -\mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \rho_{t-1}(\tau) \sqrt{r_{k,n}} \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}}, \tag{3.5}$$

which coincides with (3.4) when we set $\rho_{t-1}(\tau) = q_\alpha$. To this end, we note that $\rho_{t-1}(\tau)$ does not depend on the portfolio weights $\mathbf{w}$. As a result, finding the optimal portfolio by optimizing (3.5) is the same problem as finding the optimal portfolio based on (3.4).

4. Portfolio optimization

We now turn to the theory related to portfolio optimization. First, we describe some existing results related to frequentist and Bayesian portfolio optimization. This is followed by theory related to the Bayesian mean-VaR and mean-CVaR optimization problem which relies on some of the previous findings. Special focus is put on the global minimum VaR (GMVaR) and global minimum CVaR (GMCVaR) portfolios. The section ends with a derivation of the efficient frontiers in the mean-VaR and mean-CVaR spaces.

4.1. Portfolio optimization problems and existing solutions

The mean-variance optimization problem of Markowitz [28] and its solution provides a foundation to practical asset allocation. It was originally formulated in terms of the true population parameters of the asset return distribution, more precisely the mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$. As a result, the solution of Markowitz's optimization problem depends on these unknown quantities which have to be estimated before the implementation in practice.

The population expected portfolio return of the portfolio with weights $\mathbf{w}$ and its population variance are given by $R_P(\mathbf{w}) = \mathbf{w}^\top \boldsymbol{\mu}_0$ and $V_P(\mathbf{w}) = \mathbf{w}^\top \boldsymbol{\Sigma}_0 \mathbf{w}$, respectively. Using these notations, Markowitz's optimization problem is given by

$$\min_{\mathbf{w}:\, R_P(\mathbf{w})=R_0,\, \mathbf{w}^\top \mathbf{1}=1} V_P(\mathbf{w}), \tag{4.1}$$

where $\mathbf{1}$ denotes the k-dimensional vector of ones. The solution of (4.1) is expressed as

$$\mathbf{w}_{MV} = \mathbf{w}_{GMV} + \frac{R_0 - R_{GMV}}{s} \mathbf{M} \boldsymbol{\mu}_0 \quad \text{with} \quad \mathbf{M} = \boldsymbol{\Sigma}_0^{-1} - \frac{\boldsymbol{\Sigma}_0^{-1} \mathbf{1}\mathbf{1}^\top \boldsymbol{\Sigma}_0^{-1}}{\mathbf{1}^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{1}}, \tag{4.2}$$

where

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}_0^{-1} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{1}} \tag{4.3}$$

are the weights of the global minimum variance (GMV) portfolio whose population expected return and population variance are given by

$$R_{GMV} = \frac{\mathbf{1}^\top \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_0}{\mathbf{1}^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{1}} \quad \text{and} \quad V_{GMV} = \frac{1}{\mathbf{1}^\top \boldsymbol{\Sigma}_0^{-1} \mathbf{1}}. \tag{4.4}$$

The quantity s is the slope parameter of the mean-variance efficient frontier, the set of all optimal portfolios in the mean-variance space, and it is given by

$$s = \boldsymbol{\mu}_0^\top \mathbf{M} \boldsymbol{\mu}_0. \tag{4.5}$$

The efficient frontier itself is a parabola given by (see, e.g., Bodnar and Schmid [14], Kan and Smith [24], Merton [29])

$$(R - R_{GMV})^2 = s(V - V_{GMV}). \tag{4.6}$$

Recently, Makowitz's optimization problem has been reformulated from the Bayesian perspectives by Bauder et al. [8]. Using the portfolio expected return $R_{t-1}(\mathbf{w}) = \mathrm{E}[X_{P,t}|\mathbf{x}_{(t-1)}]$ and the portfolio variance $V_{t-1}(\mathbf{w}) = \mathrm{Var}(X_{P,t}|\mathbf{x}_{(t-1)})$ computed from the posterior predictive distribution as in (2.10), Markowitz's optimization problem from the Bayesian point of view is given by

$$\min_{\mathbf{w}:\, R_{t-1}(\mathbf{w})=R_0,\, \mathbf{w}^\top\mathbf{1}=1} V_{t-1}(\mathbf{w}), \tag{4.7}$$

with the solution expressed as

$$\mathbf{w}_{MV,t-1} = \mathbf{w}_{GMV,t-1} + \frac{R_0 - R_{GMV,t-1}}{s_{t-1}}\mathbf{M}_{t-1}\bar{\mathbf{x}}_{t-1} \quad \text{with} \quad \mathbf{M}_{t-1} = \mathbf{S}_{t-1}^{-1} - \frac{\mathbf{S}_{t-1}^{-1}\mathbf{1}\mathbf{1}^\top\mathbf{S}_{t-1}^{-1}}{\mathbf{1}^\top\mathbf{S}_{t-1}^{-1}\mathbf{1}}, \tag{4.8}$$

where

$$\mathbf{w}_{GMV,t-1} = \frac{\mathbf{S}_{t-1}^{-1}\mathbf{1}}{\mathbf{1}^\top\mathbf{S}_{t-1}^{-1}\mathbf{1}}, \quad R_{GMV,t-1} = \frac{\mathbf{1}^\top\mathbf{S}_{t-1}^{-1}\bar{\mathbf{x}}_{t-1}}{\mathbf{1}^\top\mathbf{S}_{t-1}^{-1}\mathbf{1}}, \quad \text{and} \quad V_{GMV,t-1} = \frac{d_{k,n}r_{k,n}}{d_{k,n}-2}\frac{1}{\mathbf{1}^\top\mathbf{S}_{t-1}^{-1}\mathbf{1}}. \tag{4.9}$$

The quantity s_{t-1} is one of the factors determining the slope of the Bayesian efficient frontier in the mean-variance space and it is given by

$$s_{t-1} = \bar{\mathbf{x}}_{t-1}^\top\mathbf{M}_{t-1}\bar{\mathbf{x}}_{t-1}. \tag{4.10}$$

Also from the Bayesian perspective, the efficient frontier is a parabola expressed as (see, Bauder et al. [8])

$$(R - R_{GMV,t-1})^2 = \frac{d_{k,n}-2}{d_{k,n}r_{k,n}}s_{t-1}(V - V_{GMV,t-1}). \tag{4.11}$$

In contrast to the population optimal portfolios and the efficient frontier, the Bayesian optimal portfolio and the Bayesian efficient frontier are presented in terms of the historical data that are observable up to time $t-1$, when the optimal portfolio for the next period is constructed. This constitutes the main advantage of the Bayesian approach, namely it takes the parameter uncertainty into account before the optimal portfolio is constructed.

Assuming asset returns to be normally distributed, Alexander and Baptista [1,2] extended Markowitz's optimization problem by replacing the population portfolio variance in (4.1) with the population VaR or CVaR given by

$$Q_P(\mathbf{w}) = -\mathbf{w}^\top\boldsymbol{\mu}_0 + q_{P;\alpha}\sqrt{\mathbf{w}^\top\boldsymbol{\Sigma}_0\mathbf{w}}, \tag{4.12}$$

where $q_{P;\alpha} = z_\alpha$ for VaR and $q_{P;\alpha} = \dfrac{\exp\left(-z_\alpha^2/2\right)}{(1-\alpha)\sqrt{2\pi}}$ for CVaR where z_α denotes the α-quantile of the standard normal distribution.

The mean-VaR or mean-CVaR optimization problem of Alexander and Baptista [1,2] is given by

$$\min_{\mathbf{w}:\, R_P(\mathbf{w})=R_0,\, \mathbf{w}^\top\mathbf{1}=1} Q_P(\mathbf{w}). \tag{4.13}$$

If the constraint on the expected return is omitted in (4.13), then the solution of (4.13) are the weights of the population optimal portfolios with the smallest values of VaR or CVaR at level α given by (see, Bodnar et al. [15])

$$\mathbf{w}_{GMQ} = \mathbf{w}_{GMV} + \frac{\sqrt{V_{GMV}}}{\sqrt{q_{P;\alpha}^2 - s}}\mathbf{M}\boldsymbol{\mu}_0. \tag{4.14}$$

Similarly to the mean-variance portfolio, the weights (4.14) of the population minimum VaR or CVaR portfolio cannot be computed. First, the unknown population parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ should be estimated by using historical data of asset returns and, then, the estimator of $\mathbf{w}_{GMQ}$ is constructed as a proxy of the true portfolio weights. This two-step procedure of constructing an optimal portfolio usually leads to sub-optimal solutions since the parameter uncertainty is ignored in its construction. In the next subsection, we deal with the problem from the viewpoint of Bayesian statistics which allows us to incorporate the parameter uncertainty directly into the decision process before the optimization problem is solved.

4.2. Bayesian mean-VaR and mean-CVaR optimal portfolios

The extension of the Alexander and Baptista [1,2] mean-VaR or mean-CVaR optimization problem (4.13) to the Bayesian perspectives is given by

$$\min_{\mathbf{w}:\, R_{t-1}(\mathbf{w})=R_0,\, \mathbf{w}^\top\mathbf{1}=1} Q_{t-1}(\mathbf{w}), \tag{4.15}$$

where $R_{t-1}(\mathbf{w})$ and $Q_{t-1}(\mathbf{w})$ are computed by using the posterior predictive distribution as discussed in Sections 2.1 and 3.1.

The solution of the optimization problem (4.15) can be presented in the following way

$$\underset{\mathbf{w}:\, R_{t-1}(\mathbf{w})=R_0,\, \mathbf{w}^\top \mathbf{1}=1}{\operatorname{argmin}} Q_{t-1}(\mathbf{w}) = \underset{\mathbf{w}:\, R_{t-1}(\mathbf{w})=R_0,\, \mathbf{w}^\top \mathbf{1}=1}{\operatorname{argmin}} -R_{t-1}(\mathbf{w}) + q_\alpha \sqrt{\frac{d_{k,n}-2}{d_{k,n}}}\sqrt{V_{t-1}(\mathbf{w})} = \underset{\mathbf{w}:\, R_{t-1}(\mathbf{w})=R_0,\, \mathbf{w}^\top \mathbf{1}=1}{\operatorname{argmin}} V_{t-1}(\mathbf{w}),$$

provided that $d_{k,n} > 2$. Hence, on the one hand, all solutions of the mean-VaR or mean-CVaR optimization problem (4.15) are also the solutions of the mean-variance optimization problem (4.7) and belong to the efficient frontier (4.11). On the other hand, all four optimization problems (4.1), (4.7), (4.13), and (4.15) possess a solution only if R_0 is properly chosen. For example, Eqs. (4.1) and (4.13) have solutions if and only if $R_0 > R_{GMV}$ and $R_0 > R_{GMV,t-1}$, respectively, while for solving (4.13) one requires that

$$q_{P;\alpha}^2 - s > 0. \tag{4.16}$$

Below in Theorem 4.1, we formulate the conditions needed for the existence of the Bayesian optimal portfolio in the sense of minimizing $Q_{t-1}(\mathbf{w})$.

Note that the conditions for a solution to exist in the case of the population optimization problems (4.1) and (4.13) depend on the unknown population parameters of the data generating process, and thus cannot be validated in practice. In contrast, the Bayesian formulation of the optimization problems makes it possible to specify the existence conditions in terms of the previously observed data $\mathbf{x}_{(t-1)}$ and, thus, to check them before the optimization problem is solved. Finally, the conditions on the existence of the solutions in the mean-VaR or mean-CVaR optimization problems (4.13) and (4.15) depend on the chosen level α, although the solutions themselves are independent of it.

Similarly to the mean-variance optimization problems, in order to determine under which conditions imposed on R_0 the solution of the mean-VaR or mean-CVaR optimization problem exists, one has to find the optimal portfolio with the smallest possible value of the objective function $Q_{t-1}(\mathbf{w})$, that is when the constraint $R_{t-1}(\mathbf{w}) = R_0$ is dropped from the optimization problem (4.15). The expected return of this GMVaR or GMCVaR portfolio will provide the smallest possible value for which the optimization problem (4.15) possesses a solution. To this end, we note that this is also the portfolio which a completely risk averse investor may be interested in. The following theorem expresses the variance and return of such a portfolio.

Theorem 4.1. *Let $d_{k,n} > 2$. Then, under the conditions of Proposition 2.1, the GMVaR or GMCVaR optimal portfolio exists if and only if*

$$q_\alpha^2 > r_{k,n}^{-1} s_{t-1}, \tag{4.17}$$

where s_{t-1} is defined in (4.10). Moreover, its posterior predictive expected return and variance are given by

$$R_{GMQ,t-1} = R_{GMV,t-1} + \frac{r_{k,n}^{-1} s_{t-1}}{\sqrt{q_\alpha^2 - r_{k,n}^{-1} s_{t-1}}}\sqrt{\frac{d_{k,n}-2}{d_{k,n}}}\sqrt{V_{GMV,t-1}}, \tag{4.18}$$

and

$$V_{GMQ,t-1} = \frac{q_\alpha^2}{q_\alpha^2 - r_{k,n}^{-1} s_{t-1}} V_{GMV,t-1}, \tag{4.19}$$

where $V_{GMV,t-1}$ and $R_{GMV,t-1}$ are given in (4.9).

The statement of Theorem 4.1 is proved in the appendix. Its results determine the lower bound for possible values of R_0 that can be used in the optimization problem (4.15). Since $R_{GMQ,t-1} > R_{GMV,t-1}$, we get that the set of optimal portfolios which solve (4.15) does not coincide with the set of the Bayesian mean-variance optimal portfolios, which lie on the upper part of the efficient frontier given by the parabola (4.11) in the mean-variance space.

The findings of Theorem 4.1 lead to the expression of GMVaR or GMCVaR for the selected level α expressed as

$$Q_{GMQ,t-1} = -R_{GMQ,t-1} + q_\alpha \sqrt{\frac{d_{k,n}-2}{d_{k,n}}}\sqrt{V_{GMQ,t-1}}. \tag{4.20}$$

Finally, the weights of the GMVaR or GMCVaR portfolio are deduced from the findings of Theorem 4.1 and they are presented in Theorem 4.2.

Theorem 4.2. *Let $d_{k,n} > 2$ and the inequality (4.17) holds. Then, under the conditions of Proposition 2.1, the weights of the GMVaR or GMCVaR portfolio are given by*

$$\mathbf{w}_{GMQ,t-1} = \mathbf{w}_{GMV,t-1} + \frac{r_{k,n}^{-1}\sqrt{V_{GMV,t-1}}}{\sqrt{q_\alpha^2 - r_{k,n}^{-1} s_{t-1}}}\sqrt{\frac{d_{k,n}-2}{d_{k,n}}}\mathbf{M}_{t-1}\bar{\mathbf{x}}_{t-1}. \tag{4.21}$$

4.3. Bayesian efficient frontiers in mean-VaR and mean-CVaR spaces

Earlier in this section, we proved that the solutions of the mean-VaR or mean-CVaR portfolio optimization problem (4.15) belong the Bayesian efficient frontier (4.11) in the mean-variance space. We now characterise the locations of the Bayesian mean-VaR and mean-CVaR optimal portfolios in the mean-VaR and mean-CVaR spaces, respectively. It has to be noted that the population mean-VaR efficient frontier was illustrated by Alexander and Baptista [1] under the assumption that the asset returns were multivariate normally distributed. We extend these findings in Theorem 4.3 whose proof is given in the appendix.

Theorem 4.3. *Let $d_{k,n} > 2$ and $s_{t-1} > 0$. Then, under the conditions of Proposition 2.1, the Bayesian efficient frontier in the mean-VaR or mean-CVaR space is a hyperbola given by*

$$R^2 - 2RR_{GMV,t-1} + R^2_{GMV,t-1} - \frac{a_{t-1}}{b^2}(R+Q)^2 + a_{t-1}V_{GMV,t-1} = 0, \tag{4.22}$$

where

$$a_{t-1} = \frac{d_{k,n}-2}{d_{k,n}r_{k,n}}s_{t-1} \quad \text{and} \quad b = q_\alpha\sqrt{\frac{d_{k,n}-2}{d_{k,n}}}.$$

Expressions for the mean-variance efficient frontier using the Bayesian setup was derived in Bauder et al. [8]. It holds that the mean-variance efficient frontier (4.11) is a parabola in the mean-variance space and a hyperbola in the mean-standard deviation space for $s_{t-1} > 0$. These findings are in line with the results in Merton [29], where the same conclusions were drawn for the population efficient frontier. In Theorem 4.3, we prove that the efficient frontier in the mean-VaR or mean-CVaR space is also a hyperbola under the same condition $s_{t-1} > 0$. It is interesting to note that since $\mathbf{M}_{t-1}$ is positive semi-definite with $\mathbf{M}_{t-1}\mathbf{1} = \mathbf{0}$ by construction, it always holds that $s_{t-1} \geq 0$ with $s_{t-1} = 0$ only if the elements of the vector $\bar{\mathbf{x}}_{t-1}$ are all equal. Another important observation is that both efficient frontiers (4.11) and (4.22) are determined by the same set of quantities $R_{GMV,t-1}$, $V_{GMV,t-1}$, and s_{t-1} which are computed from the historical data of asset returns. Finally, we note that the equation of the efficient frontier can be rewritten as

$$Q = q_\alpha\sqrt{\frac{(R-R_{GMV,t-1})^2}{r^{-1}_{k,n}s_{t-1}} + \frac{d_{k,n}-2}{d_{k,n}}V_{GMV,t-1}} - R.$$

4.4. Assessment of the bias in the estimators of the efficient frontier

To assess the bias in the estimators of the efficient frontier, we use the notation for the true parameters of the data generating model which we denote by $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ for the mean vector and for the covariance matrix, respectively. From the Bayesian perspective, the Dirac delta prior is assigned to $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ which put a point mass at $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$. In such a case, the posterior mass is also located at $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$.

The Dirac delta prior is treated as the true state of knowledge, while the Jeffreys prior (2.2) and the conjugate prior (2.3) reflects the investor intuition about the model parameters. As such, the derived equations of the efficient frontier are considered as generic procedures and their stochastic properties will be investigated under the true model whose parameters are endowed with the Dirac delta prior. Similarly, we also treated the conventional estimator of the efficient frontier which uses the sample mean vector and sample covariance matrix to estimate the model parameters.

The analytical expression of the population efficient frontier in the mean-VaR or mean-CVaR space is obtained by employing the Dirac delta prior and it is given by

$$R^2 - 2RR_{GMV} + R^2_{GMV} - \frac{s}{q^2_\alpha}(R+Q)^2 + sV_{GMV} = 0, \tag{4.23}$$

which is also a hyperbola in the mean-VaR or mean-CVaR space. It is fully determined by the same set of constants R_{GMV}, V_{GMV}, and s as the population efficient frontier (4.6), which is also a hyperbola in the mean standard-deviation space. To this end, the derived expression of the efficient frontier complements the findings of Alexander and Baptista [1] who presents this frontier in the empirical study without deriving its closed-form expression.

To assess the bias in the Bayesian estimators of the efficient frontier derived under the Jeffreys prior, we define

$$H_{t-1;B}(R,Q) = R^2 - 2RR_{GMV,t-1} + R^2_{GMV,t-1} - \frac{a_{t-1}}{b^2}(R+Q)^2 + a_{t-1}V_{GMV,t-1}. \tag{4.24}$$

Similarly, for the conventional sample estimator, we consider

$$H_{t-1;S}(R,Q) = R^2 - 2R\hat{R}_{GMV} + \hat{R}^2_{GMV} - \frac{\hat{s}}{q^2_\alpha}(R+Q)^2 + \hat{s}\hat{V}_{GMV}, \tag{4.25}$$

which is obtained from (4.23) by replacing the population quantities R_{GMV}, V_{GMV}, and s by their sample counterparts given by

$$\hat{R}_{GMV} = R_{GMV,t-1}, \quad \hat{V}_{GMV} = \frac{d_{k,n}-2}{d_{k,n}r_{k,n}(n-1)}V_{GMV,t-1}, \quad \text{and} \quad \hat{s} = (n-1)s_{t-1}, \tag{4.26}$$

where $R_{GMV,t-1}$, $V_{GMV,t-1}$, and s_{t-1} are obtained under the Jeffreys prior.

Due to the definition of the population efficient frontier (4.23), it should hold that the expectations $\mathrm{E}[H_{t-1;B}(R,Q)]$ and $\mathrm{E}[H_{t-1;S}(R,Q)]$ are zero for all R and Q when the Bayesian and the sample estimators of the efficient frontier are unbiased. In Theorem 4.4 we present the bias of the three estimators of the efficient frontier. The proof of Theorem 4.4 is given in the appendix.

Theorem 4.4. *Let asset returns $\mathbf{X}_1, \mathbf{X}_2, \dots$ be i.i.d. with $\mathbf{X}_i \sim N_k(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. Then*

1. *for the conventional sample of the efficient frontier, it holds for $n > k$ that*

$$\mathrm{E}[H_{t-1;S}(R,Q)] = -\frac{1}{q_\alpha^2}\left(\frac{k}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)(R+Q)^2 + \left(\frac{2}{n-k-1}s + \frac{nk-k^2+k-2}{n(n-k-1)}\right)V_{GMV}; \tag{4.27}$$

2. *under the Jeffreys prior, it holds for $n > k$ that*

$$\mathrm{E}[H_{t-1;B}(R,Q)] = -\frac{1}{q_\alpha^2}\left(\frac{k+1}{(n+1)(n-k-1)}s + \frac{(n-k)(k-1)}{(n+1)(n-k-1)}\right)(R+Q)^2 + \left(\frac{2}{n-k-1}s + \frac{nk-k^2+k-2}{n(n-k-1)}\right)V_{GMV}. \tag{4.28}$$

Both the sample estimator of the efficient frontier and the Bayesian estimator with the Jeffres prior are biased estimators of the efficient frontier. Although the bias of both estimators vanishes asymptotically when the number of assets in the portfolio k is relatively small to the sample size n, it is no longer true in more realistic situations where large-dimensional portfolios are constructed. The latter case is known in the literature as the high-dimensional asymptotic regime and the asymptotic analysis is usually carried out by assuming that $k/n \to c$ as $n \to \infty$ where c is called the concentration ratio (cf, Bai and Silverstein [4]). In the high-dimensional asymptotic regime, the asymptotic bias of the sample estimator becomes

$$\mathrm{E}[H_{t-1;S}(R,Q)] = -\frac{1}{q_\alpha^2}\frac{c}{1-c}(s+1)(R+Q)^2 + cV_{GMV}$$

$k/n \to c$ as $n \to \infty$, while for the Bayesian estimator with the Jeffreys prior it is given by

$$\mathrm{E}[H_{t-1;B}(R,Q)] = -\frac{1}{q_\alpha^2}c(R+Q)^2 + cV_{GMV}$$

$k/n \to c$ as $n \to \infty$. Although both estimators of the efficient frontier are biased in the high-dimensional setting, the bias of the sample estimator explodes when the number of assets in the portfolio approaches the sample size.

5. Simulation study

In the following section, we analyse how the Bayesian approaches compare to the conventional plug-in method via simulations. We do so by studying VaR prediction using the global minimum VaR (GMVaR) portfolio. Moreover, we analyse the different estimation methods by illustrating their corresponding efficient frontiers. When doing this comparison we also include the global minimum variance (GMV) portfolio in the analysis to see where it is located in the mean-VaR space. It should be noted that the comparison of the global minimum CVaR (GMCVaR) portfolio could be carried out similarly by evaluating VaR at many different levels.

5.1. Setup of simulation study

Throughout the simulation study, the asset returns are generated from a multivariate normal distribution. In order to not restrict the analysis to certain parameters, the mean vector and covariance matrix are randomized in each new simulation iteration. In each iteration, the mean vector and covariance matrix are estimated from a random selection of stocks that belong to the S&P 500 index using their sample counterparts at some random date between January 2010 and December 2020 (see, Table 5 in Appendix B for the list of stocks). We also consider different sample sizes and portfolio sizes by using $n \in \{100, 200\}$ and letting $k \in \{5, 10, 15, 20\}$ when $n = 100$ and $k \in \{10, 20, 30, 40\}$ when $n = 200$. Moreover, we use $\alpha \in \{0.95, 0.99\}$ to study the impact of the VaR level in the GMVaR computations. For each parameter setup, we consider 10,000 independent simulation runs when studying the accuracy of the GMVaR portfolios. We then aggregate the obtained results in all simulation runs.

Since the true parameters of the asset return distribution are known during simulation, it is possible to make comparisons with the population GMVaR portfolios as well as the population efficient frontier. The population GMVaR portfolios are constructed from the same mathematical formulas as when using the conventional method but they are based on the true

Table 1
Relative VaR exceedance frequencies for the population GMVaR portfolio and its three estimates based on simulated returns.

Parameter setup			GMVaR portfolio			
α	n	k	Jeffreys	Conjugate	Conventional	Population
0.95	100	5	0.0564	0.0601	0.0624	0.0506
		10	0.0617	0.0697	0.0755	0.0492
		15	0.0711	0.0849	0.0947	0.0510
		20	0.0859	0.1026	0.1172	0.0546
	200	10	0.0573	0.0596	0.0615	0.0512
		20	0.0692	0.0752	0.0803	0.0498
		30	0.0767	0.0876	0.0970	0.0494
		40	0.0876	0.1012	0.1172	0.0522
0.99	100	5	0.0136	0.0157	0.0165	0.0117
		10	0.0148	0.0178	0.0201	0.0105
		15	0.0180	0.0228	0.0280	0.0094
		20	0.0235	0.0331	0.0418	0.0098
	200	10	0.0113	0.0125	0.0135	0.0101
		20	0.0159	0.0185	0.0208	0.0111
		30	0.0203	0.0251	0.0299	0.0119
		40	0.0220	0.0296	0.0397	0.0112

parameters. Hence the population portfolios can be used as benchmarks for the corresponding Bayesian and conventional portfolios which are all based on parameter estimates. Similarly, the population efficient frontier can be used as a reference for the estimated efficient frontiers.

It should also be mentioned that the hyperparameters $\boldsymbol{m_0}$ and $\boldsymbol{S_0}$ of the conjugate prior are determined by using the empirical Bayesian approach (see, e.g., Bauder et al. [7]) and we set $d_0 = r_0 = n$.

5.2. Accuracy of the GMVaR portfolio

We use two measures to analyze the accuracy of the GMVaR portfolios. The first measure is the relative frequency of times the estimated VaR is exceeded, i.e,

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{-X_{\mathrm{GMVaR},i} \geq \widehat{\mathrm{VaR}}_\alpha(X_{\mathrm{GMVaR},i})\},$$

where N is the number of simulation runs, $\mathbf{1}$ is the indicator function, $X_{\mathrm{GMVaR},i}$ is the actual return of the estimated GMVaR portfolio for simulation i and $\widehat{\mathrm{VaR}}_\alpha(X_{\mathrm{GMVaR},i})$ is its predicted VaR. The latter two are calculated using Eqs. (4.20) and (4.21) in the Bayesian cases and (4.12) and (4.14) in the conventional and population cases. By the definition of VaR, an exceedance rate close to $1-\alpha$ means a good prediction of the VaR.

In Table 1, we observe that the relative VaR exceedance of the population GMVaR portfolios are always close to the target level. The only source of noise is from the number of simulation runs. For the three estimation methods we observe that the Jeffreys prior gives consistently best results followed by the conjugate prior. Compared to the conventional method, the Jeffreys prior shows an improvement between 8 and 63 percent relative to the target level when $\alpha = 0.95$.[1] Similarly, the relative improvement is between 22 and 183 percent when $\alpha = 0.99$. It also appears that the relative advantage of the Jeffreys prior becomes more apparent as k increases. However, although the Bayesian methods outperform the conventional one, they still slightly underestimate the risk. This is indicated by a relative exceedance frequency higher than $1-\alpha$ in all cases. We also note that it seems to be more difficult to correctly estimate GMVaR when $\alpha = 0.99$ compared to when $\alpha = 0.95$.

The relative VaR exceedance frequency measures how accurate the GMVaR estimates are by its definition. It does not measure how far away the estimated GMVaR is from the population GMVaR. To investigate this behaviour of our predictions we use a second measure which is the average absolute deviation of the estimated GMVaR to its population value, i.e.,

$$\frac{1}{N}\sum_{i=1}^{N}|\widehat{\mathrm{VaR}}_\alpha(X_{\mathrm{GMVaR},i}) - \mathrm{VaR}_\alpha(X_{\mathrm{GMVaR},i})|,$$

where $\mathrm{VaR}_\alpha(X_{\mathrm{GMVaR},i})$ is the VaR of the population GMVaR portfolio. Since the population GMVaR portfolio is based on the true parameter values, the VaR of the population GMVaR portfolio coincides with the true value of the minimum VaR. Hence, it can be used as a benchmark and the average absolute deviation should ideally be close to zero.

Table 2 shows the results of the GMVaR portfolio comparison using the average absolute deviation as a performance measure. As for Table 1, we observe that the Bayesian approach based on the Jeffreys prior is the best performing method

[1] The relative improvement is calculated as the difference between the conventional relative exceedance frequency and the Jeffreys relative exceedance frequency divided by the target relative exceedance frequency.

Table 2
Average absolute deviation of the VaR of the estimated GMVaR portfolios to the VaR of the population GMVaR portfolio based on simulated returns. Values inside brackets represent standard deviations.

Parameter setup			GMVaR portfolio		
α	n	k	Jeffreys	Conjugate	Conventional
0.95	100	5	0.0026 (0.0022)	0.0026 (0.0022)	0.0027 (0.0022)
		10	0.0023 (0.0018)	0.0024 (0.0020)	0.0027 (0.0021)
		15	0.0021 (0.0017)	0.0024 (0.0019)	0.0030 (0.0022)
		20	0.0020 (0.0016)	0.0026 (0.0020)	0.0036 (0.0023)
	200	10	0.0016 (0.0012)	0.0016 (0.0013)	0.0017 (0.0013)
		20	0.0014 (0.0011)	0.0016 (0.0012)	0.0020 (0.0014)
		30	0.0013 (0.0010)	0.0018 (0.0013)	0.0025 (0.0015)
		40	0.0013 (0.0010)	0.0020 (0.0013)	0.0031 (0.0015)
0.99	100	5	0.0034 (0.0028)	0.0033 (0.0027)	0.0034 (0.0028)
		10	0.0028 (0.0024)	0.0029 (0.0024)	0.0033 (0.0026)
		15	0.0026 (0.0021)	0.0028 (0.0022)	0.0037 (0.0026)
		20	0.0025 (0.0020)	0.0029 (0.0023)	0.0043 (0.0028)
	200	10	0.0020 (0.0015)	0.0020 (0.0016)	0.0021 (0.0016)
		20	0.0017 (0.0013)	0.0019 (0.0015)	0.0024 (0.0017)
		30	0.0016 (0.0013)	0.0020 (0.0015)	0.0030 (0.0018)
		40	0.0015 (0.0012)	0.0023 (0.0015)	0.0037 (0.0018)

followed by the conjugate prior. The Jeffreys prior shows generally the smallest deviations, although the different methods are quite close in their accuracy. Moreover, it seems like the Jeffreys prior is less sensitive to an increased portfolio size. This is indicated by stable results when α and n are fixed and k varies. We also note that the accuracy of all methods seem to be sightly worse for $\alpha = 0.99$ compared to $\alpha = 0.95$.

To conclude, the suggested Bayesian methods perform better at estimating the GMVaR than the conventional approach. Using the Jeffreys prior seems to give the best results, although using the conjugate prior is also beneficial compared to the conventional method. However, all of the methods are slightly underestimating GMVaR, especially for large portfolios. Moreover, it seems to be more difficult to estimate GMVaR correctly for a higher VaR level. These points are further studied in the next section where we investigate the influence of parameter uncertainty on the estimation of the whole mean-VaR efficient frontier.

5.3. Comparison of efficient frontiers

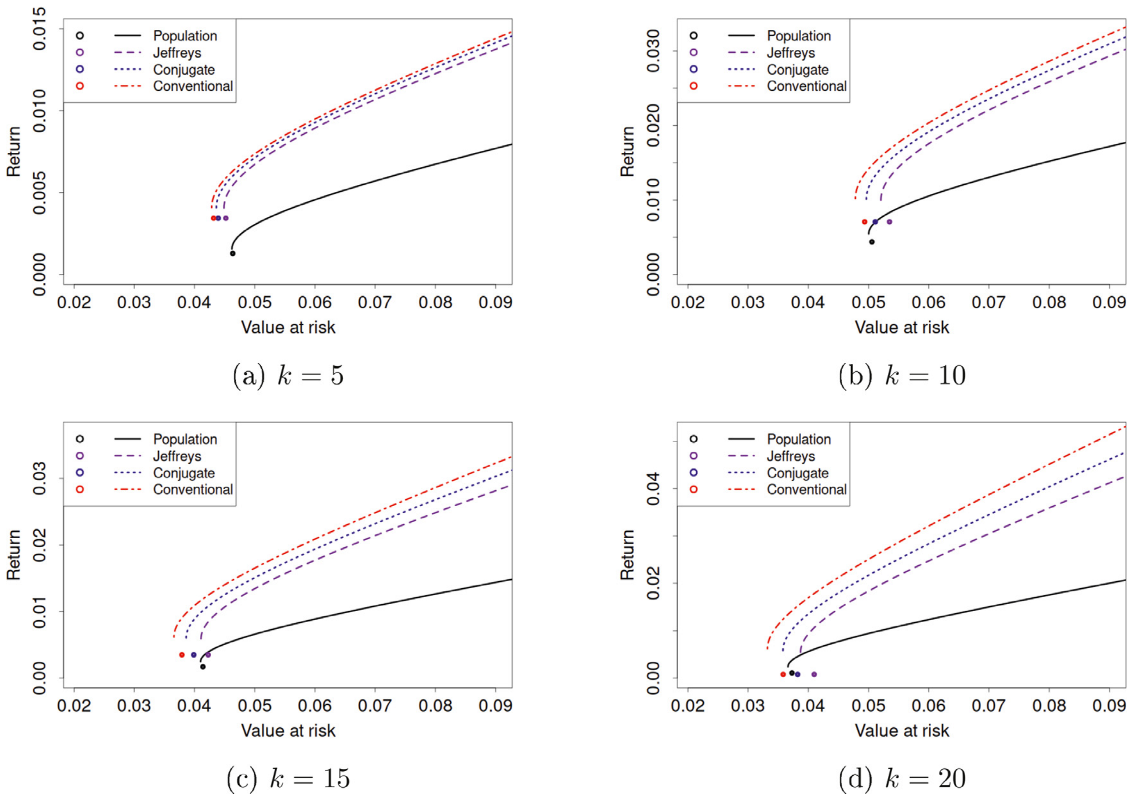
In order to get a better understanding of the impact of parameter uncertainty in mean-VaR portfolio optimization, we use the theoretical findings of Section 4.3 and plot the population mean-VaR efficient frontier together with its three estimates in Figs. 2–4. The estimates of the mean-VaR efficient frontier are computed for a single simulation run as described at the beginning of this section by using (4.22) and (4.23) for the Bayesian and conventional estimates, respectively. It should be noted that the figures present the most common results which are also observed for other simulation runs. All of the figures also show where the portfolio which globally minimizes the variance is located in the mean-VaR space using each of the methods.

The mean-VaR efficient frontiers and the locations of the GMV portfolios are depicted in Fig. 2 for $\alpha = 0.99$, $n = 100$, and $k \in \{5, 10, 15, 20\}$. We observe that all methods overestimate the location of the true efficient frontier in the mean-VaR space. Such a behaviour is similar to the one previously documented for the Markowitz efficient frontier in the mean-variance space by Bauder et al. [6], Bodnar and Bodnar [11], Siegel and Woodgate [33] among others. Namely, estimation tend to lead to overoptimistic investment opportunities where the investors expect more return for the same level of risk than the population efficient frontier determines. This becomes particularly pronounced for larger portfolios, e.g., when $k = 20$.

Among the three estimates, the Bayesian methods, particularly the one based on the Jeffereys prior, lead to the curves that are closest to the population frontier for all considered portfolio sizes. We also observe the positive effect of portfolio diversification in Fig. 2. Increasing the portfolio size typically leads to the reduction of the VaR of the GMVaR portfolio. Also, we note the positive effect on the slope parameter of the efficient frontier which usually becomes larger.

Fig. 2 also illustrates that the portfolios that minimize the variance are not located on the mean-VaR efficient frontiers. This is an expected but important observation which illustrates that an investor who is mean-variance efficient may not always be mean-VaR efficient.

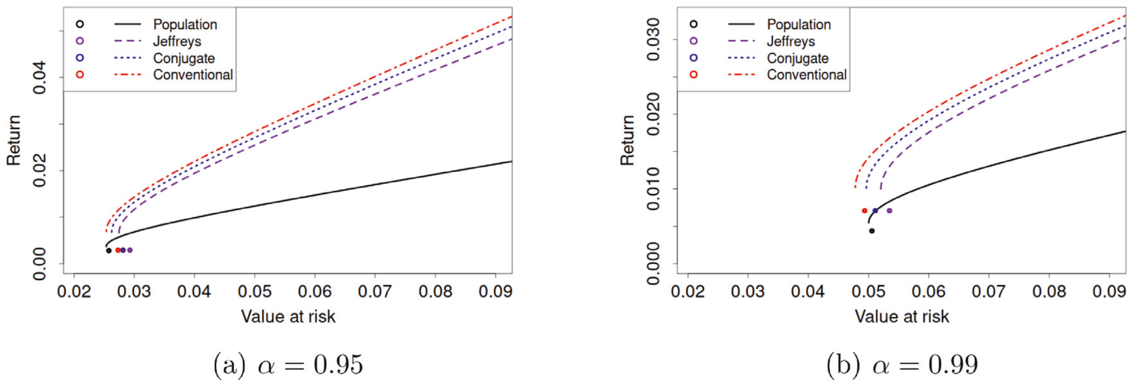
Figs. 3 and 4 demonstrate that the conclusions drawn from the results of Fig. 2 are also valid for other values of α and n. In both figures, the Bayesian approach with the Jeffreys prior provides the best fit of the population efficient frontier followed by the Bayesian estimate based on the conjugate prior. Also, we observe that the increase of the portfolio size with the simultaneous increase of the sample size leads the reduction of the VaR of the GMVaR portfolio. Moreover, the GMV portfolios are again shown to not be mean-VaR efficient.



(a) $k = 5$ (b) $k = 10$

(c) $k = 15$ (d) $k = 20$

Fig. 2. Population mean-VaR efficient frontier together with its three estimates for $n = 100$, $\alpha = 0.99$ and $k \in \{5, 10, 15, 20\}$ based on simulated returns. The locations of the GMV portfolios are marked by circles. Different scales are used on the axes for presentation purposes.



(a) $\alpha = 0.95$ (b) $\alpha = 0.99$

Fig. 3. Population mean-VaR efficient frontier together with its three estimates for $n = 100$, $k = 10$ and $\alpha \in \{0.95, 0.99\}$ based on simulated returns. The locations of the GMV portfolios are marked by circles. Different scales are used on the axes for presentation purposes.

6. Empirical illustration

We now continue the comparison between the Bayesian and conventional methodologies through an application on actual market data. As an analogy to the simulation study, we first study the accuracy of the GMVaR portfolio and investigate the behaviour of the efficient frontiers using real stock returns. Once again, we consider the cases $n \in \{100, 200\}$, $k \in \{5, 10, 15, 20\}$ when $n = 100$ and $k \in \{10, 20, 30, 40\}$ when $n = 200$, and $\alpha \in \{0.95, 0.99\}$. Next we also apply our methodology on a portfolio aggregation level by finding optimal combinations of portfolios that have been put together based on some common characteristics. We again study the VaR exceedance frequency for the GMVaR portfolio and study the resulting efficient frontiers on this aggregation level.

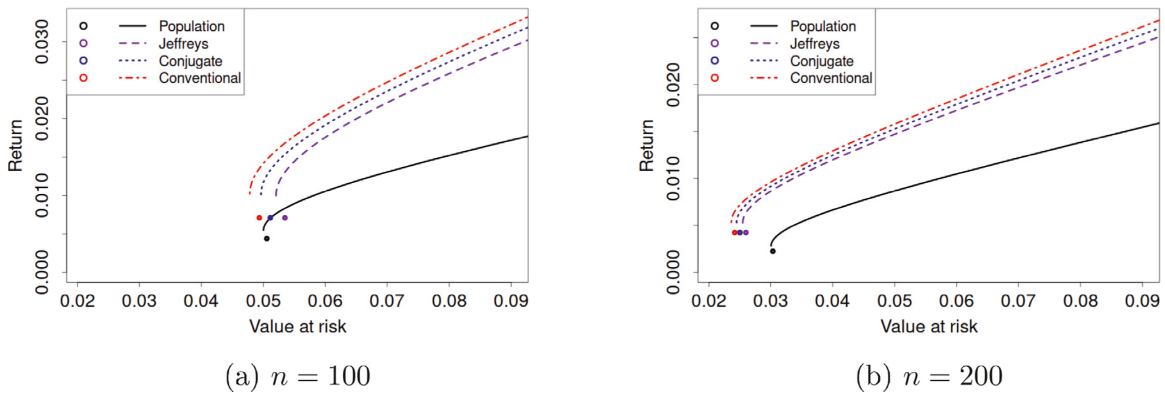


Fig. 4. Population mean-VaR efficient frontier together with its three estimates for $\alpha = 0.99$, $n \in \{100, 200\}$ and $k/n = 0.1$ based on simulated returns. The locations of the GMV portfolios are marked by circles. Different scales are used on the axes for presentation purposes.

Table 3
Relative VaR exceedance frequencies between January 2010 and December 2020 for the three estimates of the GMVaR portfolio based on weekly returns from stocks.

Parameter setup			GMVaR portfolio		
α	n	k	Jeffreys	Conjugate	Conventional
0.95	100	5	0.0662	0.0695	0.0711
		10	0.0750	0.0812	0.0859
		15	0.0865	0.0958	0.1035
		20	0.0973	0.1092	0.1221
	200	10	0.0721	0.0749	0.0769
		20	0.0831	0.0883	0.0927
		30	0.0931	0.1008	0.1091
		40	0.1022	0.1151	0.1279
0.99	100	5	0.0308	0.0324	0.0335
		10	0.0347	0.0379	0.0409
		15	0.0391	0.0449	0.0495
		20	0.0436	0.0519	0.0603
	200	10	0.0351	0.0366	0.0378
		20	0.0398	0.0436	0.0460
		30	0.0425	0.0474	0.0532
		40	0.0488	0.0562	0.0636

6.1. Description of stock market data

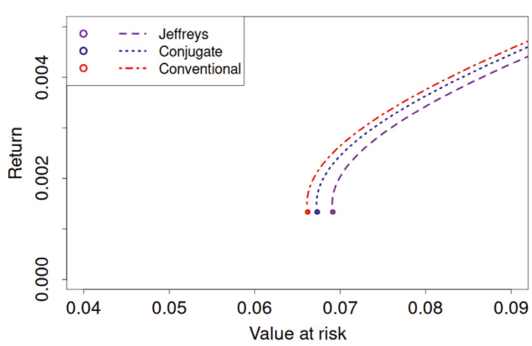
In the first part of the empirical application, we use weekly returns from stocks included in the S&P 500 index between January 2010 and December 2020 to evaluate the performance of the GMVaR portfolio and study the efficient frontiers. Weekly data of asset returns is chosen since it provides a good compromise between fulfilling the assumption of conditional normality used in the derivation of theoretical results and availability of data, i.e., the time intervals between observations are not too long.

In order to circumvent the possible bias of selecting stocks which outperform or underperform the rest of the market, we consider all stocks included in the S&P 500 by 2020 that were already part of the index by our chosen start date. The lack of public information makes it difficult to know exactly when a certain stock was added to this index, but based on [37] we have chosen to consider 215 stocks that were present in the index before the 1st of January, 2010. A complete list of the stocks is provided in Table 5 in Appendix B. We randomly choose 100 portfolios of size k from the list of stocks for each possible value of $\{n, k, \alpha\}$. Once the stocks have been selected, they are kept for the whole time period, but the weights of the GMVaR portfolio are re-calculated each week and the performance is evaluated on a weekly basis and then averaged across all sampled portfolios.

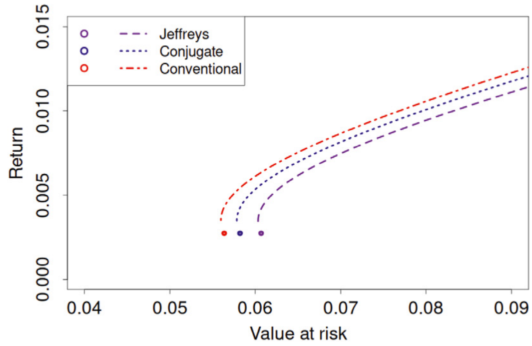
The hyperparameters when using the conjugate prior are specified as in the simulation study, i.e., by employing the empirical Bayesian approach and setting $d_0 = r_0 = n$.

6.2. Results using stock returns

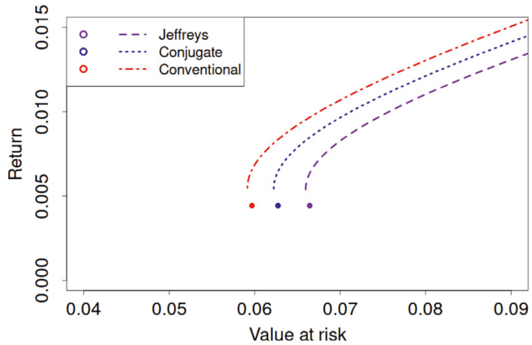
As in the simulation study, we consider the relative VaR exceedance frequency when evaluating the accuracy of the GMVaR portfolio. The result is summarized in Table 3.



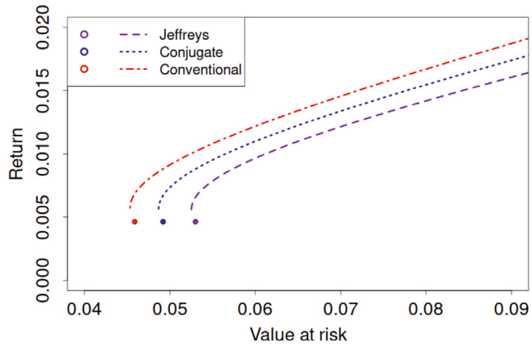
(a) $k = 5$



(b) $k = 10$



(c) $k = 15$



(d) $k = 20$

Fig. 5. Bayesian and conventional mean-VaR efficient frontiers based on stock data for $n = 100$, $\alpha = 0.99$ and $k \in \{5, 10, 15, 20\}$. The locations of the GMV portfolios are marked by circles. Different scales are used on the axes for presentation purposes.

As can be seen, the Bayesian approaches are clearly outperforming the conventional method by having exceedance frequencies closer to $1 - \alpha$. Using the Jeffreys prior gives the best results, followed by the conjugate prior, in all situations that we consider. The relative improvement of using the Jeffreys prior compared to the conventional method is between 10 and 51 percent when $\alpha = 0.95$ and between 27 and 167 percent when $\alpha = 0.99$. However, as in the simulation study, all of the methods are slightly underestimating VaR since the relative exceedance frequency is always higher than $1 - \alpha$. This is especially pronounced when k is large.

Similar results to those observed in Table 3 are also present in Fig. 5 where we plot the estimated mean-VaR efficient frontier at the end of 2020 using $\alpha = 0.99$, $n = 100$, and $k \in \{5, 10, 15, 20\}$. Both Bayesian efficient frontiers are always located under the conventional efficient frontier. While the three efficient frontiers almost coincide when $k = 5$, the difference between the conventional and Bayesian approaches becomes more clear when k becomes larger. Moreover, we again see that the GMV portfolios are not mean-VaR efficient. All of this is in line with the observations made for Fig. 2 in the simulation study. Varying α and n using the empirical data will also result in the same relationships between the efficient frontiers as shown in Figs. 3 and 4 in the simulation study, indicating the considerable overoptimism present in the construction of the conventional efficient frontier.

6.3. Portfolios based on common characteristics

In the second part of the empirical application we consider returns from equally weighted portfolios based on some common characteristics instead of individual stock returns. In the financial industry, it is common to consider such kind of grouping of assets. For instance, the Nobel price awarded model by Fama and French uses the size and book to market ratio of companies to categorise stocks [18]. It has been observed in practise that stocks that belong to the same category in the Fama and French model have more similar expected returns for a certain level of risk. Hence, by forming several equally weighted portfolios based on some factors and then consider the portfolios, instead of the individual stocks, as assets in the optimization problem, we can find an optimal combination of portfolios where the assets possibly behave more similarly. Moreover, this approach makes it easier to invest in a large amount of stocks without requiring the same amount of data as when investing in them individually.

It is typical in the Fama and French model to consider 2 levels for the size and 3 levels for the book to market ratio in the categorization, resulting in 6 possible categories, or equally weighted portfolios [18]. Another commonly used categorisation

Table 4
Relative VaR exceedance frequencies between January 2010 and December 2020 for the three estimates of the GMVaR portfolio based on returns from categorised portfolios.

Parameter setup			GMVaR portfolio		
α	n	Categories	Jeffreys	Conjugate	Conventional
0.95	100	6	0.0803	0.0846	0.0846
		25	0.1036	0.1184	0.1438
	200	6	0.0912	0.0912	0.0912
		25	0.1019	0.1100	0.1126
0.99	100	6	0.0317	0.0317	0.0317
		25	0.0528	0.0634	0.0803
	200	6	0.0402	0.0429	0.0429
		25	0.0482	0.0590	0.0617

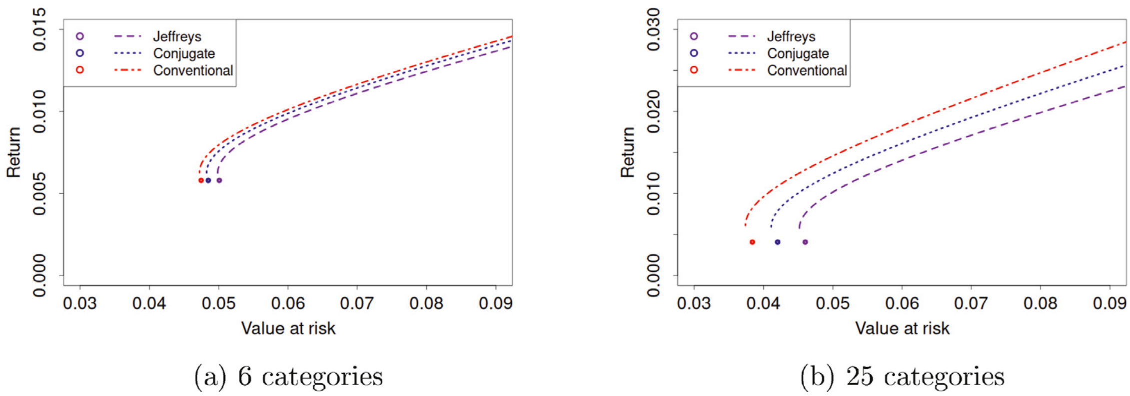


Fig. 6. Bayesian and conventional mean-VaR efficient frontiers based on data from 6 and 25 equally weighted portfolios that have been constructed based on size and book to market ratio and using $n = 100$ and $\alpha = 0.99$. The locations of the GMV portfolios are marked by circles. Different scales are used on the axes for presentation purposes.

uses 5 levels for both the size and book to market ratio which means that 25 categories, or equally weighted portfolios, are constructed. We consider weekly returns from such equally weighted portfolios based on 6 and 25 categories between January 2010 and December 2020. Each portfolio consists of possibly hundreds of stocks belonging to the NYSE, AMEX or NASDAQ. The GMVaR exceedance frequencies are studied as well as the efficient frontiers. Data for this analysis is collected from French [19] where also more detailed information about the portfolios is provided. As when analysing individual stock returns, we consider the cases $n \in \{100, 200\}$ and $\alpha \in \{0.95, 0.99\}$. The hyperparameters when using the conjugate prior are specified as previously, i.e., by using the empirical Bayesian approach and setting $d_0 = r_0 = n$.

6.4. Results using portfolio returns

Table 4 shows the relative VaR exceedance frequency of the GMVaR portfolio when the assets are equally weighted portfolios based on size and book to market ratio instead of individual stocks.

We conclude from Table 4 that the VaR exceedance frequenccies are similar to what was observed for the stock returns, although the GMVaR estimates seem to be maybe a slightly bit more optimistic. The methods based on the Bayesian approach provides more accurate estimates of the GMVaR compared to the conventional method but the differences seem to be very small when 6 categories are considered. This is in line with previous observations that the Bayesian approach has a more significant advantage when considering many assets. It is interesting to note that increasing the rolling window size n does not necessarily result in better estimates of the GMVaR portfolio, which was always the case when considering stock returns. This could be because the composition of the equally weighted portfolios change over time and hence more detailed information about the stocks in the portfolios could resolve this issue.

Fig. 6 shows mean-VaR efficient frontiers constructed from categorised portfolios at the end of 2020 when using $n = 100$ and $\alpha = 0.99$. As was the case for the stock returns, we observe that efficient frontiers corresponding to the Bayesian methods are located under the conventional efficient frontier. Moreover, by comparing Figs. 5 and 6, we note that efficient frontiers constructed by combining portfolios seem to have a smaller minimum VaR and also a higher level of return for a certain level of VaR compared to corresponding efficient frontiers constructed from stocks. This can also be confirmed numerically by constructing 100 efficient frontiers where each efficient frontier corresponds to a random selection of 6 or 25 stocks. Measured in absolute terms, it holds that the minimum VaR obtained by combining portfolios instead of stocks gives on average about 0.02 lower minimum VaR when 6 categories are considered and about 0.004 lower minimum VaR when 25 categories are used, no matter the estimation method. Moreover, the expected return of the GMVaR portfolio tends to be almost 0.003 larger when considering 6 categories and almost 0.002 larger when considering 25 categories, using any of the estimation methods. Keep in mind that the accuracy of the GMVaR estimates are quite similar when the assets are either portfolios or stocks when using $n = 100$ and $\alpha = 0.99$, indicating a possible real advantage of investing in portfolios based on common characteristics instead of individual stocks. It should be emphasized though that these results regarding the efficient frontiers are based on its appearance the last trading day of 2020, although similar results are observed also for other days.

7. Conclusion

The traditional mean-variance analysis has been a paramount foundation for the development of portfolio analysis based on one-sided risk measures which are popular in financial mathematics. However, the conventional approaches usually ignore the parameter uncertainty in the construction of an optimal portfolio. It is common to define optimal portfolios by a two-step procedure where first an optimization problem is solved and then the optimal portfolios are estimated by replacing the unknown quantities in the solutions by the corresponding sample counterparts.

The Bayesian methodology differs fundamentally from the conventional approaches with respect to what is being optimized. Investors care about the complete future risk in taking a position, not the conditional risk of having a certain position. The Bayesian framework uses the posterior predictive distribution in the optimization problem to automatically consider the parameter uncertainty and cope with the problem of future risk. This is in contrast to the conventional method which simply ignores this kind of uncertainty.

We contribute to the existing literature by formulating and solving the mean-VaR and mean-CVaR portfolio allocation problems from the perspective of Bayesian statistics. We show that mean-VaR and mean-CVaR optimal portfolios will be mean-variance efficient in our Bayesian setup, but the reverse is not always true. This similarity is due to the derived symmetrical aspects of the posterior predictive distribution, but it by no means make the analysis less relevant. It is well known that low frequency returns (for instance weekly or monthly) exhibit approximately this characteristic and long-term investors that are bounded by the Basel regulations should wish to be efficient from the VaR and CVaR perspectives and not just with respect to the portfolio variance. We have presented conditions for when this holds. Results of the simulation study and of the empirical application leads to the conclusion that the Bayesian approaches to portfolio construction provide a good alternative to the conventional procedures. In particular, the Bayesian approaches outperform the conventional one in terms of accuracy of the GMVaR portfolio and estimation of the mean-VaR efficient frontier.

Although using the Jeffreys prior gave the best results in our study, a more careful calibration of the hyperparameters of the conjugate prior could have made that one more beneficial. Moreover, even tough the Bayesian approaches reduce the underestimation of VaR considerably and shrink the estimates of the efficient frontier, they still show some overoptimism when the portfolio size is large in comparison to the sample size. Further research in this direction is needed which might lead to interesting results completing the existing findings in the direction of large-dimensional portfolio construction (see, e.g., Bodnar et al. [12], Hautsch et al. [22]).

Acknowledgements

The authors would like to thank Professor Theodore Simos and an anonymous reviewer for their helpful suggestions. This research was partly supported by the Swedish Research Council (VR) via the project “Bayesian Analysis of Optimal Portfolios and Their Risk Measures”.

Appendix A. Proofs of theoretical results

In order to get the stochastic representations presented below we also need the following result.

Lemma A.1. *Let a random variable z possess the following stochastic representation*

$$z \stackrel{d}{=} \frac{\tau_1}{\sqrt{\nu d}} + \sqrt{1+\frac{\tau_1^2}{d}}\frac{\tau_2}{\sqrt{d+1}}, \tag{A.1}$$

where $d > 0$, τ_1 and τ_2 are independent with $\tau_1 \sim t(d)$ and $\tau_2 \sim t(d+1)$. Then, z follows a t-distribution with d degrees of freedom, location parameter 0, and scale parameter $\sqrt{(\nu+1)/\nu d}$.

Proof of Lemma A1. Since τ_1 and τ_2 are independent with $\tau_2 \sim t(d+1)$, the conditional distribution of z given τ_1 is a t-distribution with $d+1$ degrees of freedom, location parameter $\tau_1/\sqrt{\nu d}$ and scale parameter $g(\tau_1)/\sqrt{d+1}$ with $g(\tau_1) = \sqrt{1+\tau_1^2/d}$. Thus the joint distribution of z and τ_1 is given by

$$f(z,\tau_1) = f(z|\tau_1)f(\tau_1)$$
$$= \frac{\Gamma\left(\frac{d+2}{2}\right)}{\Gamma\left(\frac{d+1}{2}\right)}\frac{1}{\sqrt{\pi}}\frac{1}{g(\tau_1)}\left(1+\left(\frac{z-\frac{\tau_1}{\sqrt{\nu d}}}{g(\tau_1)}\right)^2\right)^{-\frac{d+2}{2}}\frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\frac{1}{\sqrt{\pi d}}\left(1+\frac{\tau_1^2}{d}\right)^{-\frac{d+1}{2}}$$

$$\propto \frac{1}{g(\tau_1)}\left(1+\left(\frac{z-\frac{\tau_1}{\sqrt{\nu d}}}{g(\tau_1)}\right)^2\right)^{-\frac{d+2}{2}} g(\tau_1)^{-(d+1)} = \left(g(\tau_1)^2+\left(z-\frac{\tau_1}{\sqrt{\nu d}}\right)^2\right)^{-\frac{d+2}{2}}$$
$$= \left(1+\frac{1}{d}[z,\tau_1]\begin{bmatrix} d & -\frac{\sqrt{\nu d}}{\nu} \\ -\frac{\sqrt{\nu d}}{\nu} & \frac{\nu+1}{\nu}\end{bmatrix}\begin{bmatrix} z \\ \tau_1\end{bmatrix}\right)^{-\frac{d+2}{2}}.$$

The last expression is the kernel of a multivariate t-distribution with d degrees of freedom, location vector $\boldsymbol{\nu}=\mathbf{0}$ and dispersion matrix $\boldsymbol{\Omega}$ given by

$$\boldsymbol{\Omega} = \begin{bmatrix} \frac{\nu+1}{\nu d} & \frac{\sqrt{\nu d}}{\nu d} \\ \frac{\sqrt{\nu d}}{\nu d} & 1 \end{bmatrix}.$$

Hence, the marginal distribution of z is also a t-distribution with d degrees of freedom, location 0 and scale $\sqrt{(\nu+1)/\nu d}$ (see, e.g., p. 15 in Kotz and Nadarajah [27]). $\square$

Proof of Proposition 2.1. Bauder et al. [8] characterized the posterior predictive distribution of the portfolio return by deriving the stochastic representation of $\widehat{X}_{P,t}$ given by

$$\widehat{X}_{P,t} \stackrel{d}{=} m+\sqrt{s}\left(\frac{\tau_1}{\sqrt{\nu d}}+\sqrt{1+\frac{\tau_1^2}{d}}\frac{\tau_2}{\sqrt{d+1}}\right)$$

with $m=\mathbf{w}^\top\bar{\mathbf{x}}_{t-1,J}$, $s=\mathbf{w}^\top\mathbf{S}_{t-1,J}\mathbf{w}$, $\nu=n$, and $d=n-k$ under the Jeffreys prior and with $m=\mathbf{w}^\top\bar{\mathbf{x}}_{t-1,I}$, $s=\mathbf{w}^\top\mathbf{S}_{t-1,I}\mathbf{w}$, $\nu=n+r_0$ and $d=n+d_0-2k$ under the conjugate prior. The application of Lemma A.1 leads to the statement of the proposition. $\square$

Proof of Proposition 3.1. The statement of the proposition follows from the fact that $\mathbf{w}^\top\bar{\mathbf{x}}_{t-1}$ is linear in $\mathbf{w}$ and that, since $\mathbf{S}_{t-1}$ is positive definite, $\sqrt{\mathbf{w}^\top\mathbf{S}_{t-1}\mathbf{w}}$ can be regarded as the Euclidean norm of $\mathbf{S}_{t-1}^{1/2}\mathbf{w}$ where $\mathbf{S}_{t-1}^{1/2}$ is the symmetric square root of $\mathbf{S}_{t-1}$. Since $q_\alpha>0$ and the Euclidean norm is convex, the result follows. $\square$

Proof of Theorem 4.1. Let $c_{k,n}=\frac{d_{k,n}r_{k,n}}{d_{k,n}-2}$. Since the solution of

$$\min_{\mathbf{w}:\,\mathbf{w}^\top\mathbf{1}=1} Q_{t-1}(\mathbf{w}),$$

belongs to the Bayesian efficient frontier (4.11) in the mean-variance space, it can be found by solving the univariate optimization problem given by

$$\min_{V:V\geq V_{GMV,t-1}} -R_{GMV,t-1}-c_{k,n}^{-1/2}\sqrt{s_{t-1}}\sqrt{V-V_{GMV,t-1}}+q_\alpha c_{k,n}^{-1/2}\sqrt{r_{k,n}}\sqrt{V} \tag{A.2}$$

where $R_{GMV,t-1}$ and $V_{GMV,t-1}$ are given in (4.9) and s_{t-1} is defined in (4.11). The solution of (A.2) solves

$$q_\alpha\sqrt{r_{k,n}}\frac{1}{\sqrt{V}} = \sqrt{s_{t-1}}\frac{1}{\sqrt{V-V_{GMV,t-1}}} \tag{A.3}$$

and it is given by

$$V_{GMQ,t-1} = \frac{q_\alpha^2}{q_\alpha^2-r_{k,n}^{-1}s_{t-1}}V_{GMV,t-1}. \tag{A.4}$$

where it obviously holds that $V_{GMQ,t-1}>V_{GMV,t-1}$ as soon as $q_\alpha^2-r_{k,n}^{-1}s_{t-1}>0$, which coincides with the second order condition needed to ensure that $V_{GMQ,t-1}$ is the solution of (A.2).

Finally, $R_{GMQ,t-1}$ is obtained from (4.11) and it is given by

$$\begin{aligned} R_{GMQ,t-1} &= R_{GMV,t-1}+\sqrt{c_{k,n}^{-1}s_{t-1}}\sqrt{\frac{q_\alpha^2}{q_\alpha^2-r_{k,n}^{-1}s_{t-1}}V_{GMV,t-1}-V_{GMV,t-1}} \\ &= R_{GMV,t-1}+\frac{r_{k,n}^{-1}s_{t-1}}{\sqrt{q_\alpha^2-r_{k,n}^{-1}s_{t-1}}}\sqrt{\frac{d_{k,n}-2}{d_{k,n}}}\sqrt{V_{GMV,t-1}}. \end{aligned}$$

$\square$

Proof of Theorem 4.3. From (4.11) and (3.4), we get

$$\frac{(R-R_{GMV,t-1})^2}{a_{t-1}}+V_{GMV,t-1}=V \quad \text{and} \quad V=\left(\frac{R+Q}{b}\right)^2 \tag{A.5}$$

where

$$a_{t-1} = \frac{d_{k,n}-2}{d_{k,n}r_{k,n}} s_{t-1} \quad \text{and} \quad b = q_\alpha \sqrt{\frac{d_{k,n}-2}{d_{k,n}}}.$$

Finally, we note that the Bayesian efficient frontier in the mean-VaR or mean-CVaR space can be rewritten as

$$R^2 - 2RR_{GMV,t-1} + R^2_{GMV,t-1} - \frac{a_{t-1}}{b^2}R^2 - 2\frac{a_{t-1}}{b^2}RQ - \frac{a_{t-1}}{b^2}Q^2 + a_{t-1}V_{GMV,t-1} = 0,$$

which is a hyperbola in the mean-VaR or mean-CVaR space for $s_{t-1} > 0$ (see, Bronshtein et al. [16, e.g., Section 3.5.2.11]) since

$$-\frac{a_{t-1}}{b^2}\left(1 - \frac{a_{t-1}}{b^2}\right) - \frac{a^2_{t-1}}{b^4} = -\frac{a_{t-1}}{b^2} = -\frac{s_{t-1}}{r_{k,n}q^2_\alpha} < 0.$$

□

Proof of Theorem 4.4.

1. From the results of Lemma A.1 of Bodnar and Schmid [14] and Lemma 2.1 of Bodnar and Bodnar [11], it holds that

$$\begin{aligned}
\mathrm{E}[\hat{R}_{GMV}] &= R_{GMV},\\
\mathrm{E}[\hat{R}^2_{GMV}] &= R^2_{GMV} + \left(\frac{n-2}{n(n-k-1)} + \frac{1}{n-k-1}s\right)V_{GMV},\\
\mathrm{E}[\hat{V}_{GMV}] &= \frac{n-k}{n-1}V_{GMV},\\
\mathrm{E}[\hat{s}] &= \frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}.
\end{aligned}$$

Using that $\hat{s}$ and $\hat{V}_{GMV}$ are independent (see Lemma 1A in Bodnar and Schmid [14]), we get

$$\begin{aligned}
\mathrm{E}[H_{t-1;S}(R,Q)] &= R^2 - 2R\mathrm{E}[\hat{R}_{GMV}] + \mathrm{E}[\hat{R}^2_{GMV}] - \frac{\mathrm{E}[\hat{s}]}{q^2_\alpha}(R+Q)^2 + \mathrm{E}[\hat{s}]\mathrm{E}[\hat{V}_{GMV}]\\
&= R^2 - 2RR_{GMV} + R^2_{GMV} + \left(\frac{n-2}{n(n-k-1)} + \frac{1}{n-k-1}s\right)V_{GMV}\\
&\quad -\frac{1}{q^2_\alpha}\left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)(R+Q)^2 + \left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)\frac{n-k}{n-1}V_{GMV}\\
&= \frac{s}{q^2_\alpha}(R+Q)^2 - sV_{GMV} + \left(\frac{n-2}{n(n-k-1)} + \frac{1}{n-k-1}s\right)V_{GMV}\\
&\quad -\frac{1}{q^2_\alpha}\left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)(R+Q)^2 + \left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)\frac{n-k}{n-1}V_{GMV}\\
&= -\frac{1}{q^2_\alpha}\left(\frac{k}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)(R+Q)^2 + \left(\frac{2}{n-k-1}s + \frac{nk-k^2+k-2}{n(n-k-1)}\right)V_{GMV}.
\end{aligned}$$

2. Similarly, for the Bayesian efficient frontier derived by employing the Jeffreys prior we get using (4.24) and (4.26) that

$$\begin{aligned}
\mathrm{E}[H_{t-1;B}(R,Q)] &= R^2 - 2R\mathrm{E}[\hat{R}_{GMV}] + \mathrm{E}[\hat{R}^2_{GMV}] - \frac{\mathrm{E}[a_{t-1}]}{b^2}(R+Q)^2 + \mathrm{E}[a_{t-1}]\mathrm{E}[V_{GMV,t-1}]\\
&= R^2 - 2RR_{GMV} + R^2_{GMV} + \left(\frac{n-2}{n(n-k-1)} + \frac{1}{n-k-1}s\right)V_{GMV}\\
&\quad -\frac{1}{q^2_\alpha}\frac{n(n-k)}{(n+1)(n-1)}\left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)(R+Q)^2\\
&\quad +\left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)\frac{n-k}{n-1}V_{GMV}\\
&= \frac{s}{q^2_\alpha}(R+Q)^2 - sV_{GMV} + \left(\frac{n-2}{n(n-k-1)} + \frac{1}{n-k-1}s\right)V_{GMV}\\
&\quad -\frac{1}{q^2_\alpha}\left(\frac{n(n-k)}{(n+1)(n-k-1)}s + \frac{(n-k)(k-1)}{(n+1)(n-k-1)}\right)(R+Q)^2\\
&\quad +\left(\frac{n-1}{n-k-1}s + \frac{(n-1)(k-1)}{n(n-k-1)}\right)\frac{n-k}{n-1}V_{GMV}
\end{aligned}$$

$$= -\frac{1}{q_\alpha^2}\left(\frac{k+1}{(n+1)(n-k-1)}s + \frac{(n-k)(k-1)}{(n+1)(n-k-1)}\right)(R+Q)^2$$
$$+\left(\frac{2}{n-k-1}s + \frac{nk-k^2+k-2}{n(n-k-1)}\right)V_{GMV}.$$

□

Appendix B. List of stocks

Table 5 presents the list of stocks considered for the parameter estimation in Sections 5 and 6.

Table 5
Stocks considered in the numerical applications.

MMM	ABT	ADBE	AES	AFL	A	APD	AKAM	ALL
GOOG	AMZN	AEE	AXP	AIG	AMT	AMP	ABC	AMGN
APH	ADI	ANTM	AON	APA	AIV	AAPL	AMAT	ADM
AIZ	T	ADSK	ADP	AZO	AVB	AVY	BLL	BAC
BK	BAX	BDX	BBY	BIIB	BKNG	BXP	BSX	BF.B
CHRW	COG	COF	CAH	CCL	CBRE	CNP	CF	SCHW
CI	CINF	CTAS	CSCO	C	CTXS	CLX	CME	CMS
CTSH	CMA	CAG	STZ	COST	CSX	CMI	DVA	XRAY
DVN	DFS	DOV	DUK	EMN	ECL	EA	EMR	EOG
EFX	EQR	EL	EXPE	EXPD	FAST	FDX	FIS	FISV
FLIR	FLS	FMC	GPS	GIS	GPC	GILD	GL	GS
GWW	HRB	HAS	PEAK	HES	HD	HON	HRL	HST
HPQ	ITW	INTC	ICE	IPG	IFF	INTU	ISRG	IVZ
IRM	J	SJM	JNJ	JPM	JNPR	KEY	KIM	LB
LH	LEN	LLY	LNC	LIN	LMT	LOW	MRO	MMC
MAS	MA	MCD	MDT	MCHP	MU	MSFT	TAP	NDAQ
NOV	NTAP	NWL	NEM	NEE	NKE	JWN	NOC	NUE
NVDA	ORLY	OXY	ORCL	PCAR	PH	PBCT	PKI	PM
PXD	PNC	PFG	PGR	PLD	PRU	PSA	PHM	PWR
DGX	RL	RF	RSG	RHI	ROP	ROST	CRM	SLB
SHW	SPG	SNA	LUV	SWK	SYK	SYY	TGT	TXT
TJX	TRV	TFC	UNH	UPS	UNM	VFC	VAR	VTR
VRSN	VZ	V	VMC	WMT	WBA	DIS	WEC	WFC
WELL	WDC	WMB	WYNN	XLNX	YUM	ZBH	ZION	

References

[1] G.J. Alexander, A.M. Baptista, Economic implications of using a mean-VaR model for portfolio selection: a comparison with mean-variance analysis, J. Econ. Dyn. Control 26 (7) (2002) 1159–1193.
[2] G.J. Alexander, A.M. Baptista, A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model, Manag. Sci. 50 (9) (2004) 1261–1273.
[3] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, Coherent measures of risk, Math. Finance 9 (3) (1999) 203–228.
[4] Z. Bai, J.W. Silverstein, Spectral Analysis of Large Dimensional Random Matrices, Springer, New York, 2010.
[5] C.B. Barry, Portfolio analysis under uncertain means, variances, and covariances, J. Finance 29 (1974) 515–522.
[6] D. Bauder, R. Bodnar, T. Bodnar, W. Schmid, Bayesian estimation of the efficient frontier, Scand. J. Stat. 46 (2019) 802–830.
[7] D. Bauder, T. Bodnar, N. Parolya, W. Schmid, Bayesian inference of the multi-period optimal portfolio for an exponential utility, J. Multivar. Anal. 175 (2020) 104544.
[8] D. Bauder, T. Bodnar, N. Parolya, W. Schmid, Bayesian mean–variance analysis: optimal portfolio selection under parameter uncertainty, Quant. Finance 21 (2021) 221–242.
[9] J.M. Bernardo, A.F. Smith, Bayesian Theory, vol. 405, John Wiley & Sons, 2000.
[10] F. Black, R. Litterman, Global portfolio optimization, Financ. Anal. J. 48 (1992) 28–43.
[11] O. Bodnar, T. Bodnar, On the unbiased estimator of the efficient frontier, Int. J. Theor. Appl. Finance 13 (2010) 1065–1073.
[12] T. Bodnar, S. Dmytriv, N. Parolya, W. Schmid, Tests for the weights of the global minimum variance portfolio in a high-dimensional setting, IEEE Trans. Signal Process. 67 (17) (2019) 4479–4493.
[13] T. Bodnar, S. Mazur, Y. Okhrin, Bayesian estimation of the global minimum variance portfolio, Eur. J. Oper. Res. 256 (2017) 292–307.
[14] T. Bodnar, W. Schmid, Econometrical analysis of the sample efficient frontier, Eur. J. Finance 15 (3) (2009) 317–335.
[15] T. Bodnar, W. Schmid, T. Zabolotskyy, Minimum VaR and minimum CVaR optimal portfolios: estimators, confidence regions, and tests, Stat. Risk Model. Appl. Finance Insur. 29 (4) (2012) 281–313.
[16] I.N. Bronshtein, K.A. Semendyayev, G. Musiol, H. Mühlig, Handbook of Mathematics, Springer Science & Business Media, 2015.
[17] S.J. Brown, Optimal Portfolio Choice Under Uncertainty: A Bayesian Approach, University of Chicago, 1976 Ph.D. thesis.
[18] E.F. Fama, K.R. French, Common risk factors in the returns on stocks and bonds, J. Financ. Econ. 33 (1992) 3–56.
[19] K.R. French, Current research returns, 2022, [Online; accessed 3-January-2022] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
[20] P.A. Frost, J.E. Savarino, An empirical Bayes approach to efficient portfolio selection, J. Financ. Quant. Anal. 21 (3) (1986) 293–305.
[21] A.K. Gupta, D.K. Nagar, Matrix Variate Distributions, Chapman and Hall/CRC, Boca Raton, 2000.
[22] N. Hautsch, L.M. Kyj, P. Malec, Do high-frequency data improve high-dimensional portfolio allocations? J. Appl. Econom. 30 (2015) 263–290.

[23] X. Huang, H. Di, Uncertain portfolio selection with background risk, Appl. Math. Comput. 276 (2016) 284–296.
[24] R. Kan, D.R. Smith, The distribution of the sample minimum-variance frontier, Manag. Sci. 54 (7) (2008) 1364–1380.
[25] R. Klein, V. Bawa, The effect of estimation risk on optimal portfolio choice, J. Financ. Econ. 3 (1976) 215–231.
[26] P. Kolm, G. Ritter, On the Bayesian interpretation of Black–Litterman, Eur. J. Oper. Res. 258 (2) (2017) 564–572.
[27] S. Kotz, S. Nadarajah, Multivariate T-Distributions and Their Applications, Cambridge University Press, 2004.
[28] H. Markowitz, Portfolio selection, J. Finance 7 (1) (1952) 77–91.
[29] R.C. Merton, An analytic derivation of the efficient portfolio frontier, J. Financ. Quant. Anal. 7 (4) (1972) 1851–1872.
[30] R.C. Merton, On estimating the expected return on the market: an exploratory investigation, J. Financ. Econ. 8 (4) (1980) 323–361.
[31] S. Muhinyuza, T. Bodnar, M. Lindholm, A test on the location of the tangency portfolio on the set of feasible portfolios, Appl. Math. Comput. 386 (2020) 125519.
[32] L. Ortiz-Gracia, C.W. Oosterlee, Efficient VaR and expected shortfall computations for nonlinear portfolios within the delta-gamma approach, Appl. Math. Comput. 244 (2014) 16–31.
[33] A.F. Siegel, A. Woodgate, Performance of portfolios optimized with estimation error, Manag. Sci. 53 (2007) 1005–1015.
[34] Y. Simaan, The opportunity cost of mean–variance choice under estimation risk, Eur. J. Oper. Res. 234 (2) (2014) 382–391.
[35] R.F. Stambaugh, Analyzing investments whose histories differ in length, J. Financ. Econ. 45 (1997) 285–331.
[36] R.S. Tsay, Analysis of Financial Time Series, John Wiley & Sons, 2010.
[37] Wikipedia contributors, List of s&p 500 companies – wikipedia, the free encyclopedia, 2020, https://en.wikipedia.org/wiki/List_of_S&P_500_companies[Online; Accessed 14-July-2020].
[38] R.L. Winkler, Bayesian models for forecasting future security prices, J. Financ. Quant. Anal. (1973) 387–405.