

# Bayesian mean–variance analysis: optimal portfolio selection under parameter uncertainty

DAVID BAUDER<sup>†</sup>, TARAS BODNAR<sup>‡</sup>, NESTOR PAROLYA \*§ and WOLFGANG SCHMID<sup>¶</sup>

<sup>†</sup>Department of Mathematics, Humboldt University of Berlin, Berlin, Germany

<sup>‡</sup>Department of Mathematics, Stockholm University, Stockholm, Sweden

<sup>§</sup>Delft Institute of Applied Mathematics, Delft University of Technology, Delft, Netherlands

<sup>¶</sup>Department of Statistics, European University Viadrina, Frankfurt (Oder), Germany

(Received 9 January 2019; accepted 20 March 2020; published online 13 May 2020)

The paper solves the problem of optimal portfolio choice when the parameters of the asset returns distribution, for example the mean vector and the covariance matrix, are unknown and have to be estimated by using historical data on asset returns. Our new approach employs the Bayesian posterior predictive distribution which is the distribution of future realizations of asset returns given the observable sample. The parameters of posterior predictive distributions are functions of the observed data values and, consequently, the solution of the optimization problem is expressed in terms of data only and does not depend on unknown quantities. By contrast, the optimization problem of the traditional approach is based on unknown quantities which are estimated in the second step, and lead to a suboptimal solution. We also derive a very useful stochastic representation of the posterior predictive distribution whose application not only gives the solution of the considered optimization problem, but also provides the posterior predictive distribution of the optimal portfolio return which can be used to construct a prediction interval. A Bayesian efficient frontier, the set of optimal portfolios obtained by employing the posterior predictive distribution, is constructed as well. Theoretically and using real data we show that the Bayesian efficient frontier outperforms the sample efficient frontier, a common estimator of the set of optimal portfolios which is known to be overoptimistic.

**Keywords:** Optimal portfolio; Posterior predictive distribution; Parameter uncertainty; Efficient frontier; Stochastic representation; Black–Litterman model

**JEL Classification:** C11, C13, C44, C58, C63

## 1. Introduction

The fundamental goal of portfolio theory is to allocate optimally investments between different assets. Mean–variance optimization is a quantitative tool which allows one to make this allocation by considering the trade-off between the risk of a portfolio and its return. The basic concepts of modern portfolio theory were developed by Markowitz (1952) who introduced a mean–variance portfolio optimization procedure in which investors incorporate their preferences towards the risk and the expected return to seek the best allocation of wealth. This is attained by selecting portfolios that maximize expected portfolio return subject to achieving a prespecified level of risk or, equivalently, that minimize the variance subject to achieving a prespecified level of expected return. The

mean–variance analysis of Markowitz is an important tool for both practitioners and researchers in the financial sector today.

The classical problems and pitfalls of mean–variance analysis are mainly related to extreme weights that often occur when the sample efficient portfolio is constructed. This point was discussed in detail by Merton (1980) who presented an estimator of the instantaneous expected return on the market in a log-normal diffusion price model and showed its slow convergence. Moreover, it was proved that the estimates of the variances and covariances of the asset returns are more accurate than the estimates of the means. Best and Grauer (1991) argued that optimal portfolios are very sensitive to the level of expected returns. Therefore, improving the technique of mean estimation has recently become a key issue of the portfolio optimization problem. The same challenge is also present when the covariance matrix needs to be estimated. To this end, Broadie (1993) showed that the estimated efficient frontier, the set of all mean–variance optimal

\*Corresponding author. Email: [n.parolya@tudelft.nl](mailto:n.parolya@tudelft.nl)

portfolios, overestimates the expected returns of portfolios for different levels of estimation error. A similar conclusion has also been drawn in more recent studies by Basak *et al.* (2005), Siegel and Woodgate (2007) and Bodnar and Bodnar (2010).

An alternative approach to deal with the parameter uncertainty in portfolio analysis is to employ the methods of Bayesian statistics (cf. Barry 1974, Brown 1976, Klein and Bawa 1976, Frost and Savarino 1986, Aguilar and West 2000, Rachev *et al.* 2008, Avramov and Zhou 2010, Sekerke 2015, Bodnar *et al.* 2017, Bauder *et al.* 2018, 2019). It is remarkable that the Bayesian approach is potentially more attractive since: (i) it uses prior information about quantities of interest; (ii) it facilitates the use of fast, intuitive, and easily implementable numerical algorithms in order to simulate complex economic quantities; (iii) it accounts for estimation risk and model uncertainty in the portfolio choice problem. First applications of Bayesian statistics to portfolio analysis during the 1970s were completely based on noninformative or data-based priors. Bawa *et al.* (1979) provided an excellent early survey of such applications. The Bayesian approaches based on the diffusion prior are usually comparable with the classical methods of portfolio selection. However, if some of the risky assets have longer histories than others, then the Bayesian approaches under the diffuse prior lead to different results (see Stambaugh 1997). Jorion (1986) introduced the hyperparameter prior approach in the spirit of the Bayes–Stein shrinkage prior, whereas Black and Litterman (1992) defended an informal Bayesian analysis with economic arguments and equilibrium relations. They derived the Black–Litterman model which leads to more stable and more diversified portfolios than simple mean–variance optimization. Unfortunately, the application of this model requires a broad variety of data, some of which may be hard to find. Recent studies by Pástor (2000) and Pástor and Stambaugh (2000) centred prior beliefs around values implied by asset pricing theories. In particular, Pástor and Stambaugh (2000) investigated the portfolio choices of mean–variance–optimizing investors who use sample evidence to update prior beliefs centred on either the risk-based or characteristic-based pricing models. Tu and Zhou (2010) argued that the investment objective provides a useful prior for portfolio selection and proposed an optimal combination of the naive equally weighted portfolio rule with one of the four sophisticated strategies—the Markowitz (1952) rule, the Jorion (1986) rule, the MacKinlay and Pástor (2000) rule, or the Kan and Zhou (2007) rule—as a way to improve performance. Finally, Kacperczyk and Damien (2011) and Kacperczyk *et al.* (2013) discussed the application of Bayesian semi-parametric models, while Brandt and Santa-Clara (2006) considered the Bayesian approach in the multi-period optimal portfolio choice problem.

We contribute to the existing literature on optimal portfolio selection by formulating an optimization problem in terms of the posterior predictive distribution and solving it. Using the available information about the development of asset returns which is present in their historical observations, the aim is to construct an optimal portfolio by taking into account investor’s preferences. The conventional approach consists of two steps: (i) first, the optimization problem is solved with the solution depending on the unknown parameters of the

asset return distribution; (ii) second, the optimal portfolio weights, which are the solutions of the optimization problem, are estimated by applying the historical observations of the asset returns. The second step is always needed in practical applications, since the expression of optimal portfolio weights resulting from the first step usually involve the unknown population parameters of the asset return distribution. Replacing these parameters by their sample estimators leads to additional uncertainty in the decision process which has been ignored for a long time in financial literature. It is important to note that following this two-step approach, the obtained solution is only suboptimal and it can deviate considerably from the optimal (population) portfolio obtained in the first stage.

In this paper, we propose a new approach, where the solution of the investor’s optimization problem is obtained by employing the posterior predictive distribution which takes parameter uncertainty into account before the optimal portfolio choice problem is solved. As a result, its solution is presented in terms of historical data and is independent of the unknown parameters of the asset returns distribution. Consequently, it can be directly applied in practice and, in contrast to the conventional approach, it is optimal.

The rest of the paper is organized as follows. Main theoretical results are given in Section 2. Here, we characterize the posterior predictive distribution of the asset returns by developing a very helpful stochastic representation (Theorem 1). This stochastic representation provides not only a way how the future realization of portfolio returns can be simulated, but it is also used to derive the first two moments needed for the considered optimization problem. Section 2.2 deals with constructing optimal portfolios by maximizing the posterior mean–variance utility function, while the expression of the Bayesian efficient frontier is derived in Section 2.3. In Section 2.4 the optimal portfolio choice problem is solved by employing the informative conjugate prior, that can be considered as an extension of the Black–Litterman model, a popular approach in the financial literature. Section 3 presents a numerical comparison of the two Bayesian approaches between each other as well as to methods based on frequentist statistics. The theoretical results are implemented in an empirical study in Section 4, while Section 5 provides a conclusion. The technical derivations are given in an appendix.

## 2. Mean–variance analysis under parameter uncertainty

### 2.1. Posterior predictive distribution

Let  $\mathbf{X}_t$  denote the  $k$ -dimensional vector of returns on asset at time  $t$ . Assume that a sample of size  $n$  of asset returns  $\mathbf{x}_{t-n}, \dots, \mathbf{x}_{t-1}$ , realizations of  $\mathbf{X}_{t-n}, \dots, \mathbf{X}_{t-1}$ , is available which provides the information set  $\mathcal{F}_{t-1}$  and let  $\mathbf{x}_{(t-1)} = (\mathbf{x}_{t-n}, \dots, \mathbf{x}_{t-1})$  be the observation matrix at time  $t-1$ . Consequently, an investor makes a decision by optimizing preferences using information  $\mathcal{F}_{t-1}$ .

Before the decision problem is formulated in Section 2.2, we first derive the predictive posterior distribution  $p(\mathbf{X}_t | \mathbf{x}_{(t-1)})$  of  $\mathbf{X}_t$  given the previous observation of asset returns summarized in  $\mathbf{x}_{(t-1)}$ . The derivation of  $p(\mathbf{X}_t | \mathbf{x}_{(t-1)})$  is based on the methods of Bayesian statistics which provide well-established

techniques for providing inferences of future realizations of asset returns given information  $\mathcal{F}_{t-1}$ .

In the following we assume that the asset returns  $\mathbf{X}_1, \mathbf{X}_2, \dots$  are infinitely exchangeable and multivariate centred spherically symmetric (see, Bernardo and Smith 2000, Section 4.4 for the definition and properties). This assumption is very general and it implies that neither the unconditional distribution of the asset returns is normal nor that they are independently distributed. Moreover, the unconditional distribution of the asset returns appears to be heavy-tailed which is usually observed for financial data (see, e.g. Bradley and Taqqu 2003).

Parameterizing the density function of  $\mathbf{X}_{(t-1)} = (\mathbf{X}_{t-n}, \dots, \mathbf{X}_{t-1})$  by the parameter  $\boldsymbol{\theta}$ , the posterior distribution of  $\boldsymbol{\theta}$  is obtained by applying the Bayes theorem and it is given by

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

where  $\pi(\boldsymbol{\theta})$  denotes the prior and  $f(\mathbf{x}_{(t-1)}|\boldsymbol{\theta})$  is the likelihood function of  $\mathbf{X}_{(t-1)}$ . The posterior distribution  $\boldsymbol{\theta}$  is then used to derive the posterior predictive distribution of the portfolio return at time  $t$  expressed as

$$X_{p,t} = \mathbf{w}^\top \mathbf{X}_t, \quad (2)$$

where  $\mathbf{w} = (w_1, \dots, w_p)^\top$  is the  $k$ -dimensional vector of portfolio weights.

The posterior distribution (1) is employed in the derivation of the posterior predictive distribution as follows:

$$f(x_{p,t}|\mathbf{x}_{(t-1)}) = \int_{\boldsymbol{\theta} \in \Theta} f(x_{p,t}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathbf{x}_{(t-1)}) d\boldsymbol{\theta}. \quad (3)$$

Due to the integration present in the definition of the posterior predictive distribution, it is possible to obtain the analytical expression of  $f(x_{p,t}|\mathbf{x}_{(t-1)})$  only in very rare cases. Moreover, the integration in (3) could also be high-dimensional, which makes the application of numerical methods very time consuming and also questions the quality of their numerical approximation. In Theorem 1, we derive a stochastic representation for the posterior predictive distribution  $f(x_{p,t}|\mathbf{x}_{(t-1)})$  which can be very easily used to draw samples from this distribution as well as to compute its expected value and variance analytically. Finally, it has to be noted that the application of the stochastic representation describing the distribution of random quantities has been used both in the frequentist statistics (see, e.g. Givens and Hoeting 2012, Gupta *et al.* 2013) and in the Bayesian statistics (cf. Bodnar *et al.* 2017).

**THEOREM 1** Let  $\mathbf{X}_1, \mathbf{X}_2, \dots$  be infinitely exchangeable and multivariate centred spherically symmetric. Let  $\pi(\boldsymbol{\theta}) = |\mathbf{F}|^{1/2}$  be Jeffreys' prior where  $|\mathbf{A}|$  denotes the determinant of a square matrix  $\mathbf{A}$  and  $\mathbf{F} = -\mathbb{E}(\partial^2 \log(f(\mathbf{x}_{(t-1)}|\boldsymbol{\theta}))/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top)$  is the Fisher information matrix. Assume  $n > k$ . Then the stochastic representation of the random variable  $\hat{X}_{p,t}$  whose density is the posterior predictive distribution (3) is given by

$$\hat{X}_{p,t} \stackrel{d}{=} \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}} \times \left( \frac{\tau_1}{\sqrt{n(n-k)}} + \sqrt{1 + \frac{\tau_1^2}{n-k} \frac{\tau_2}{\sqrt{n-k+1}}} \right),$$

where

$$\bar{\mathbf{x}}_{t-1} = \frac{1}{n} \sum_{i=t-n}^{t-1} \mathbf{x}_i \quad \text{and} \quad \mathbf{S}_{t-1} = \sum_{i=t-n}^{t-1} (\mathbf{x}_i - \bar{\mathbf{x}}_{t-1})(\mathbf{x}_i - \bar{\mathbf{x}}_{t-1})^\top, \quad (4)$$

and  $\tau_1, \tau_2$  are independent with  $\tau_1 \sim t_{n-k}$  and  $\tau_2 \sim t_{n-k+1}$ . The symbol ' $\stackrel{d}{=}$ ' denotes the equality in distribution.

The proof of Theorem 1 is given in the [appendix](#). Its results provide an easy way how a random sample from the posterior distribution of  $f(x_{p,t}|\mathbf{x}_{(t-1)})$  can be simulated as summarized in Algorithm 1:

---

#### Algorithm 1

---

- (i) generate  $\tau_1^{(b)} \sim t_{n-k}$  and  $\tau_2^{(b)} \sim t_{n-k+1}$ ;
- (ii) compute

$$\hat{X}_{p,t}^{(b)} = \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}} \times \left( \frac{\tau_1^{(b)}}{\sqrt{n(n-k)}} + \sqrt{1 + \frac{(\tau_1^{(b)})^2}{n-k} \frac{\tau_2^{(b)}}{\sqrt{n-k+1}}} \right)$$

- (iii) Repeat steps (i) and (ii) for  $b = 1, \dots, B$  resulting in an independent sample  $\hat{X}_{p,t}^{(1)}, \dots, \hat{X}_{p,t}^{(B)}$  from the posterior predictive distribution (3).
- 

The generated sample  $\hat{X}_{p,t}^{(1)}, \dots, \hat{X}_{p,t}^{(B)}$  is used to calculate important characteristics of the distribution  $f(x_{p,t}|\mathbf{x}_{(t-1)})$ , like the mean, the variance, the credible interval, etc. To this end, we note that the condition  $n > k$  ensures that  $\mathbf{S}_t$  is positive definite and, hence, it is invertible.

Another important application of Theorem 1 provides us with the analytical expression of the expected value and the variance of the posterior predictive distribution  $f(x_{p,t}|\mathbf{x}_{(t-1)})$ . These findings are formulated in Corollary 1.

**COROLLARY 1** Under the conditions of Theorem 1, let  $n - k > 2$ . Then:

$$\mathbb{E}(\mathbf{w}^\top \mathbf{X}_t|\mathbf{x}_{(t-1)}) = \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} \quad (5)$$

and

$$\text{Var}(\mathbf{w}^\top \mathbf{X}_t|\mathbf{x}_{(t-1)}) = c_{k,n} \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w} \quad \text{with} \quad c_{k,n} = \frac{1}{n-k-1} + \frac{2n-k-1}{n(n-k-1)(n-k-2)}. \quad (6)$$

The proof of Corollary 1 is given in the [appendix](#). Its results are used in the next section where the expressions of optimal portfolio weights are given.

## 2.2. Mean–variance optimal portfolios

The mean–variance investor constructs an optimal portfolio at time  $t-1$  for the next period by maximizing the mean–variance utility function given by

$$\begin{aligned} U(\mathbf{w}) &= \mathbb{E}(X_{p,t}|\mathbf{x}_{(t-1)}) - \frac{\gamma}{2} \text{Var}(X_{p,t}|\mathbf{x}_{(t-1)}) \\ &= \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} - \frac{c_{k,n}\gamma}{2} \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w} \end{aligned} \quad (7)$$

under the constraint that the whole wealth is invested into the selected assets, i.e.  $\mathbf{w}^\top \mathbf{1} = 1$  where  $\mathbf{1}$  denotes the  $k$ -dimensional vector of ones. The quantity  $\gamma > 0$  stands for the coefficient of the investor's risk aversion and describes the investor's attitude towards risk.

In contrast to the conventional approach that involves the unknown parameters of the asset return distribution in its formulation, the optimization problem in (7) already incorporates the parameter uncertainty by using the available information summarized in the data matrix  $\mathbf{x}_{(t-1)}$ . As a result, the output of solving (7) is the formula for optimal portfolio weights that could be directly applied in practice, while the estimation of optimal portfolio weights is required in the conventional methods that leads to the suboptimality of the resulting portfolio.

The optimization problem in (7) is similar to the optimization problem in the conventional approach (see Ingersoll 1987, Okhrin and Schmid 2006) with the exception that the risk aversion coefficient is multiplied by the constant  $c_{k,n}$ . As a result, the solution of (7) is given by

$$\begin{aligned} \mathbf{w}_{MV,\gamma} &= \frac{\mathbf{S}_{t-1}^{-1} \mathbf{1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} + \gamma^{-1} c_{k,n}^{-1} \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1} \quad \text{with} \\ \mathbf{Q}_{t-1} &= \mathbf{S}_{t-1}^{-1} - \frac{\mathbf{S}_{t-1}^{-1} \mathbf{1} \mathbf{1}' \mathbf{S}_{t-1}^{-1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} \end{aligned} \quad (8)$$

together with the expected return and the variance expressed as

$$R_{MV,\gamma} = \frac{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \bar{\mathbf{x}}_{t-1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} + \gamma^{-1} c_{k,n}^{-1} \bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1} \quad (9)$$

and

$$V_{MV,\gamma} = \frac{c_{k,n}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} + \gamma^{-2} c_{k,n}^{-1} \bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1}, \quad (10)$$

respectively, where we use that  $\mathbf{Q}_{t-1} \mathbf{1} = \mathbf{0}$  and  $\mathbf{Q}_{t-1} \mathbf{S}_{t-1} \mathbf{Q}_{t-1} = \mathbf{Q}_{t-1}$  in (10).

Additionally to the formulae of the optimal portfolio weights, the expected return and the variance of the mean–variance optimal portfolios presented in (8)–(10), the Bayesian approach allows to characterize the posterior predictive distribution of the constructed optimal portfolio. This is achieved by applying the results of theorem 1 where the weights of an arbitrary portfolio are replaced by the optimal portfolio weights given in (8). Then, the posterior predictive distribution of the optimal portfolio return is obtained via simulations as described after theorem 1 by replacing  $\mathbf{w}$  with  $\mathbf{w}_{MV,\gamma}$  as in (8). This is a very important result which allows the whole characterization of the stochastic behaviour of optimal portfolio return and is a great advantage with respect to

the conventional approach where the point estimator is only present.

We conclude this section by noting that the original Markowitz problem (see Markowitz 1952, 1959) is solved in the same way. In the mean variance analysis of Markowitz, the optimization problem is given by: (i) minimizing the portfolio variance for a given level of the expected return  $R_0$  or (ii) maximizing the expected return for the given level of the variance  $V_0$ . In the first case the optimal portfolio weights are given by (8) with

$$\gamma^{-1} = c_{k,n} \frac{\left( R_0 - \frac{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \bar{\mathbf{x}}_{t-1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} \right)}{\bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1}}, \quad (11)$$

while in the second case the weights are obtained from (8) with

$$\gamma^{-1} = c_{k,n} \frac{\sqrt{c_{k,n}^{-1} V_0 - \frac{1}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}}}}{\sqrt{\bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1}}}. \quad (12)$$

## 2.3. Bayesian efficient frontier

Equations (9) and (10) determine the set of all optimal portfolios obtained as solutions of (7) for  $\gamma > 0$ . Solving these two equations with respect to  $\gamma$  leads to a set in the mean–variance space where all mean–variance optimal portfolios lie. We call this set the (objective) Bayesian efficient frontier since the non-informative Jeffreys prior is employed in its derivation. It is given by

$$(R - R_{GMV})^2 = \frac{\bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1}}{c_{k,n}} (V - V_{GMV}), \quad (13)$$

where

$$R_{GMV} = \frac{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \bar{\mathbf{x}}_{t-1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} \quad \text{and} \quad V_{GMV} = \frac{c_{k,n}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} \quad (14)$$

are the expected return and the variance of the global minimum variance portfolio, i.e. the mean–variance optimal portfolio with the smallest variance, whose weights are expressed as

$$\mathbf{w}_{GMV} = \frac{\mathbf{S}_{t-1}^{-1} \mathbf{1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}}. \quad (15)$$

The quantity  $s = \bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1} / c_{k,n}$  is the slope parameter of the efficient frontier which is equal to the amount of the excess squared return with respect to the return of the global minimum variance portfolio when the variance is increased by one. Finally, we note that the Bayesian efficient frontier is a parabola in the mean–variance space which is the same finding as obtained by the conventional approach (see Merton 1972).

## 2.4. Subjective Bayesian approach: extended Black–Litterman model

The results of Sections 2.1–2.3 were obtained by assigning the non-informative Jeffreys prior to the model parameter  $\mu$  and

$\Sigma$  corresponding to objective Bayesian inference in statistical literature. In this section we discuss an alternative Bayesian approach which is based on the (extended) Black–Litterman model (cf. Black and Litterman 1992). The latter model corresponds to the application of an informative prior for  $\mu$  and  $\Sigma$  and, thus, it is referred as the subjective Bayesian approach.

In order to incorporate expert knowledge in the construction of an optimal portfolio, Black and Litterman (1992) suggested to employ the normal prior for the mean vector  $\mu$ . This approach is known in financial literature as the Black–Litterman model. Below we consider an extension of this model by also including a prior on  $\Sigma$  in the decision process. To this end, it leads to the application of the informative conjugate prior for  $\mu$  and  $\Sigma$  given by

$$\mu|\Sigma \sim \mathcal{N}_k\left(\mathbf{m}_0, \frac{1}{r_0}\Sigma\right), \quad (16)$$

$$\Sigma \sim \mathcal{IW}_k(d_0, \mathbf{S}_0), \quad (17)$$

where  $\mathbf{m}_0$ ,  $r_0$ ,  $d_0$ ,  $\mathbf{S}_0$  are additional model parameters known as hyperparameters. The symbol  $\mathcal{N}_k(\mathbf{m}_0, \Sigma/r_0)$  denotes the  $k$ -dimensional normal distribution with mean vector  $\mathbf{m}_0$  and covariance matrix  $\Sigma/r_0$ , while  $\mathcal{IW}_k(d_0, \mathbf{S}_0)$  stands for the inverse Wishart distribution with  $d_0$  degrees of freedom and parameter matrix  $\mathbf{S}_0$ . The prior mean  $\mathbf{m}_0$  reflects the prior expectation about  $\mu$ , while  $\mathbf{S}_0$  determines the prior beliefs about  $\Sigma$ . The other two hyperparameters  $r_0$  and  $d_0$  are known as precision parameters for  $\mathbf{m}_0$  and  $\mathbf{S}_0$ , respectively.

In Theorem 2 we present a stochastic representation from the posterior predictive distribution of the portfolio return derived under the application of the prior (16)–(17). The proof of the theorem is presented in the [appendix](#).

**THEOREM 2** *Let  $\mathbf{X}_1, \mathbf{X}_2, \dots$  be infinitely exchangeable and multivariate centred spherically symmetric. Assume  $n + d_0 - 2k > 0$ . Then, under the application of the informative prior (16)–(17), the stochastic representation of the random variable  $\hat{X}_{p,t}$  whose density is the posterior predictive distribution (3) is given by*

$$\begin{aligned} X_{p,t} \stackrel{d}{=} & \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,I} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w}} \left( \frac{\eta_1}{\sqrt{(n+r_0)(n+d_0-2k)}} \right. \\ & \left. + \sqrt{1 + \frac{\eta_1^2}{n+d_0-2k} \frac{\eta_2}{\sqrt{n+d_0-2k+1}}} \right), \end{aligned}$$

where

$$\begin{aligned} \bar{\mathbf{x}}_{t-1,I} &= \frac{n\bar{\mathbf{x}}_{t-1} + r_0\mathbf{m}_0}{n+r_0} \quad \text{and} \quad \mathbf{S}_{t-1,I} \\ &= \mathbf{S}_{t-1} + \mathbf{S}_0 + nr_0 \frac{(\mathbf{m}_0 - \bar{\mathbf{x}}_{t-1,I})(\mathbf{m}_0 - \bar{\mathbf{x}}_{t-1,I})^\top}{n+r_0}, \end{aligned} \quad (18)$$

and  $\eta_1$  and  $\eta_2$  are independent with  $\eta_1 \sim t_{n+d_0-2k}$  and  $\eta_2 \sim t_{n+d_0-2k+1}$ .

Similarly to Theorem 1, the findings of Theorem 2 allow to simulate samples from the posterior distribution of  $f(x_{p,t}|\mathbf{x}_{(t-1)})$  in a simple way given by:

---

**Algorithm 2**


---

- (i) generate  $\eta_1^{(b)} \sim t_{n+d_0-2k}$  and  $\eta_2^{(b)} \sim t_{n+d_0-2k+1}$ ;
- (ii) compute

$$\begin{aligned} \hat{X}_{I,p,t}^{(b)} &= \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,I} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w}} \\ &\times \left( \frac{\eta_1^{(b)}}{\sqrt{(n+r_0)(n+d_0-2k)}} \right. \\ &\left. + \sqrt{1 + \frac{(\eta_1^{(b)})^2}{n+d_0-2k} \frac{\eta_2^{(b)}}{\sqrt{n+d_0-2k+1}}} \right) \end{aligned}$$

- (iii) Repeat steps (i) and (ii) for  $b = 1, \dots, B$  resulting in independent sample  $\hat{X}_{I,p,t}^{(1)}, \dots, \hat{X}_{I,p,t}^{(B)}$  from the posterior predictive distribution (3).
- 

Furthermore, the closed-form expressions of the expected value and of the variance of the posterior predictive distribution  $f(x_{p,t}|\mathbf{x}_{(t-1)})$  is computed as shown in Corollary 2 whose proof is given in the [appendix](#).

**COROLLARY 2** *Under the conditions of Theorem 2, let  $n + d_0 - 2k > 2$ . Then:*

$$\mathbb{E}(\mathbf{w}^\top \mathbf{X}_t | \mathbf{x}_{(t-1)}) = \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,I} \quad (19)$$

and

$$\mathbb{V}ar(\mathbf{w}^\top \mathbf{X}_t | \mathbf{x}_{(t-1)}) = q_{k,n} \mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w} \quad (20)$$

with

$$\begin{aligned} q_{k,n} &= \frac{1}{n+d_0-2k-1} \\ &+ \frac{2n+r_0+d_0-2k-1}{(n+r_0)(n+d_0-2k-1)(n+d_0-2k-2)}. \end{aligned}$$

Substituting (19) and (20) in (7) we find the weights of the optimal portfolios in the case of the extended Black–Litterman model expressed as

$$\begin{aligned} \mathbf{w}_{MV,\gamma} &= \frac{\mathbf{S}_{t-1,I}^{-1} \mathbf{1}}{\mathbf{1}' \mathbf{S}_{t-1,I}^{-1} \mathbf{1}} + \gamma^{-1} q_{k,n}^{-1} \mathbf{Q}_{t-1,I} \bar{\mathbf{x}}_{t-1,I} \quad \text{with} \\ \mathbf{Q}_{t-1,I} &= \mathbf{S}_{t-1,I}^{-1} - \frac{\mathbf{S}_{t-1,I}^{-1} \mathbf{1} \mathbf{1}' \mathbf{S}_{t-1,I}^{-1}}{\mathbf{1}' \mathbf{S}_{t-1,I}^{-1} \mathbf{1}} \end{aligned} \quad (21)$$

with the expected return and the variance expressed as

$$R_{MV,\gamma} = \frac{\mathbf{1}' \mathbf{S}_{t-1,I}^{-1} \bar{\mathbf{x}}_{t-1,I}}{\mathbf{1}' \mathbf{S}_{t-1,I}^{-1} \mathbf{1}} + \gamma^{-1} q_{k,n}^{-1} \bar{\mathbf{x}}_{t-1,I}^\top \mathbf{Q}_{t-1,I} \bar{\mathbf{x}}_{t-1,I} \quad (22)$$

and

$$V_{MV,\gamma} = \frac{q_{k,n}}{\mathbf{1}' \mathbf{S}_{t-1,I}^{-1} \mathbf{1}} + \gamma^{-2} q_{k,n}^{-1} \bar{\mathbf{x}}_{t-1,I}^\top \mathbf{Q}_{t-1,I} \bar{\mathbf{x}}_{t-1,I}. \quad (23)$$

Although the expression of the optimal portfolio weights obtained from the (extended) Black–Litterman model looks



similar to the one obtained in the case of the (objective) Bayesian optimal portfolio (8), they are in fact completely different due to the definition of  $\bar{\mathbf{x}}_{t-1,I}$  and  $\mathbf{S}_{t-1,I}$  in (18). In contrast to the latter approach which is based on the observed sample only, the weights resulting from the Black–Litterman model incorporate the expert knowledge about the asset returns. As a result, the Black–Litterman Bayesian optimal portfolios do not obviously belong to the efficient frontier as given in (13), but they create their own set of optimal portfolios (see (24)), which we call the Black–Litterman efficient frontier. This frontier is obtained by solving (22) and (23) with respect to  $\gamma$  resulting in a set in the mean–variance space where all mean–variance optimal portfolios lie following the (extended) Black–Litterman model. It is given by

$$(R - R_{GMV,I})^2 = \frac{\bar{\mathbf{x}}_{t-1,I}^\top \mathbf{Q}_{t-1,I} \bar{\mathbf{x}}_{t-1,I}}{q_{k,n}} (V - V_{GMV,I}), \quad (24)$$

where

$$R_{GMV,I} = \frac{\mathbf{1}^\top \mathbf{S}_{t-1,I}^{-1} \bar{\mathbf{x}}_{t-1,I}}{\mathbf{1}^\top \mathbf{S}_{t-1,I}^{-1} \mathbf{1}} \quad \text{and} \quad V_{GMV,I} = \frac{q_{k,n}}{\mathbf{1}^\top \mathbf{S}_{t-1,I}^{-1} \mathbf{1}} \quad (25)$$

are the expected return and the variance of the Black–Litterman global minimum variance portfolio, whose weights are given by

$$\mathbf{w}_{GMV,I} = \frac{\mathbf{S}_{t-1,I}^{-1} \mathbf{1}}{\mathbf{1}^\top \mathbf{S}_{t-1,I}^{-1} \mathbf{1}}. \quad (26)$$

Also, in the case of the (extended) Black–Litterman model, the efficient frontier is the parabola in the mean–variance space. However, its location also depends on the expert knowledge, that is on  $\mathbf{m}_0$  and  $\mathbf{S}_0$  as well as on the beliefs on this knowledge expressed by  $r_0$  and  $d_0$ . As a result, it might significantly deviate from the Bayesian efficient frontier given by (13). On the other side, the application of the Bernstein–von Mises theorem (cf. Bernardo and Smith 2000) ensures that as the sample size increases the differences between the two efficient frontiers (13) and (24) as well as between the optimal portfolios (8) and (21) become negligible.

### 3. Numerical study

The results of Section 2 are obtained from the viewpoint of Bayesian statistics. In this section we compare these two Bayesian approaches of the construction of optimal portfolios between each other as well as to the method based on the frequentist statistics (see, e.g. Jobson and Korkie 1981, Okhrin and Schmid 2006, Bodnar et al. 2018, 2019).

#### 3.1. Conventional approach

Let  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  be the mean vector and the covariance matrix of the asset returns. Then the traditional approach to construct an optimal portfolio consists of two steps (see, e.g. Ingersoll 1987, Okhrin and Schmid 2006):

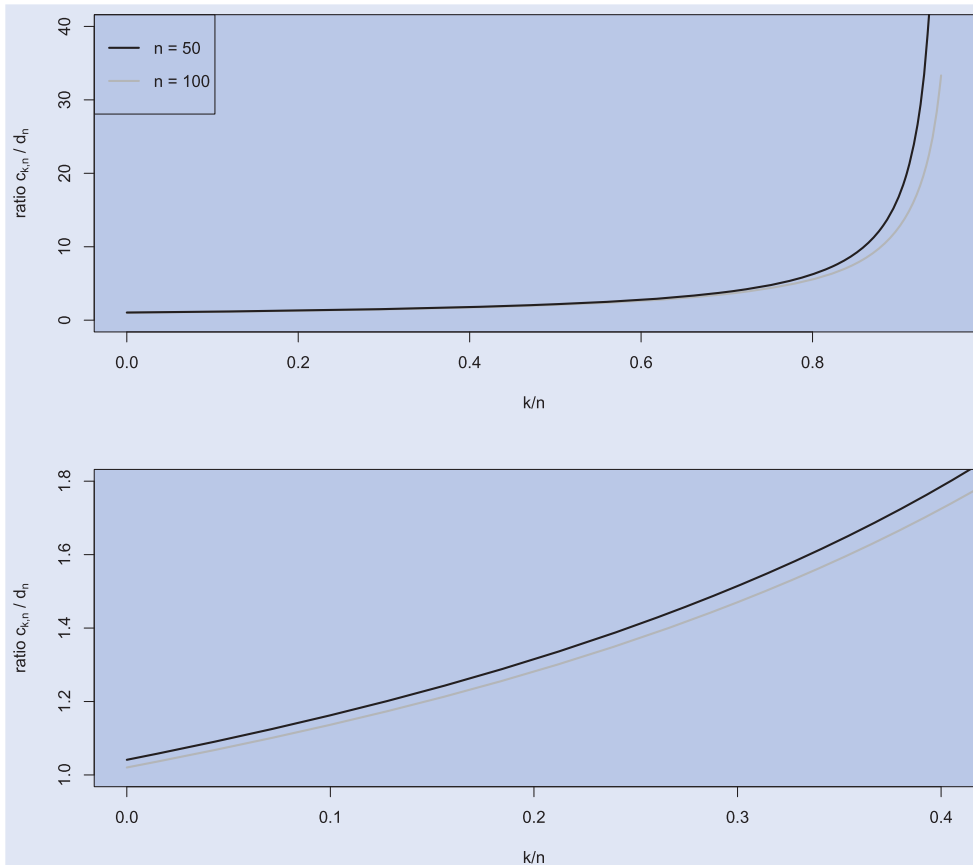


Figure 1. The ratio  $c_{k,n}/d_n$  plotted as a function of  $k/n$  for  $k/n \in [0, 0.95]$  and  $n \in \{50, 100\}$ .

Table 1. Average absolute deviation (AD) of the estimated portfolio expected return and of the estimated portfolio variance from their population values.

$\gamma = 50$		AD of portfolio expected return				AD of portfolio variance			
Bayesian Black–Litterman Sample Robust, MVE Robust, MCD		$n = 50$	$n = 75$	$n = 100$	$n = 130$	$n = 50$	$n = 75$	$n = 100$	$n = 130$
	$k = 5$	<b>0.1017</b>	<b>0.0822</b>	<b>0.0713</b>	<b>0.0621</b>	<b>0.0020</b>	<b>0.0016</b>	<b>0.0014</b>	<b>0.0012</b>
		1.0889	0.7257	0.5438	0.4140	0.0218	0.0145	0.0109	0.0083
		0.1186	0.0912	0.0768	0.0660	0.0024	0.0018	0.0015	0.0013
		0.3824	0.2284	0.1751	0.1319	0.0076	0.0046	0.0035	0.0026
		0.2494	0.1984	0.1661	0.1459	0.0050	0.0040	0.0033	0.0029
		<b>0.2541</b>	<b>0.1944</b>	<b>0.1660</b>	<b>0.1439</b>	<b>0.0051</b>	<b>0.0039</b>	<b>0.0033</b>	<b>0.0029</b>
		2.6794	1.7051	1.2657	0.9547	0.0536	0.0341	0.0253	0.0191
	$k = 10$	0.4127	0.2730	0.2183	0.1765	0.0083	0.0055	0.0044	0.0035
		1.0597	0.5949	0.4299	0.3305	0.0212	0.0119	0.0086	0.0066
		2.4460	0.9509	0.5799	0.3900	0.0489	0.0190	0.0116	0.0078
		<b>0.8878</b>	<b>0.6205</b>	<b>0.5116</b>	<b>0.4236</b>	<b>0.0178</b>	<b>0.0124</b>	<b>0.0102</b>	<b>0.0085</b>
		8.5916	4.8591	3.3871	2.4422	0.1718	0.0972	0.0677	0.0488
	$k = 25$	3.9959	1.9720	1.3382	0.9466	0.0799	0.0394	0.0268	0.0189
		11.349	5.7082	3.3371	2.1340	0.2269	0.1142	0.0667	0.0427
		25.837	9.3763	5.2220	3.2756	0.5167	0.1875	0.1044	0.0655
		<b>2.3320</b>	<b>1.2120</b>	<b>0.9153</b>	<b>0.7521</b>	<b>0.0466</b>	<b>0.0242</b>	<b>0.0183</b>	<b>0.0150</b>
		16.5132	8.1638	5.1806	3.5884	0.3303	0.1633	0.1036	0.0718
	$k = 40$	27.9857	7.4658	4.2657	2.8229	0.5597	0.1493	0.0853	0.0565
		73.6087	20.148	12.158	7.5067	1.4722	0.4030	0.2432	0.1501
		533.452	33.396	16.574	9.7837	10.669	0.6679	0.3315	0.1957

Notes: The smallest values are depicted in bold. Five considered estimators are the (objective) Bayesian estimator (first line in each panel), the estimator resulting from the (extended) Black–Litterman model (second line in each panel), the sample estimator (third line in each panel), the robust MVE estimator (fourth line in each panel), and the robust MCD estimator (fifth line in each panel) for several values of  $k \in \{5, 10, 25, 40\}$  and  $n \in \{50, 75, 100, 130\}$  with  $\gamma = 50$ . The asset returns are assumed to be conditionally normally distributed with small variances.

Table 2. Average absolute deviation (AD) of the estimated portfolio expected return and of the estimated portfolio variance from their population values.

$\gamma = 50$		AD of portfolio expected return				AD of portfolio variance			
Bayesian Black–Litterman Sample Robust, MVE Robust, MCD		$n = 50$	$n = 75$	$n = 100$	$n = 130$	$n = 50$	$n = 75$	$n = 100$	$n = 130$
	$k = 5$	<b>0.0111</b>	<b>0.0091</b>	<b>0.0079</b>	<b>0.0070</b>	<b>0.0002</b>	<b>0.0002</b>	<b>0.0002</b>	<b>0.0001</b>
		0.0932	0.0629	0.0473	0.0361	0.0018	0.0012	0.0009	0.0007
		0.0133	0.0102	0.0086	0.0075	0.0003	<b>0.0002</b>	<b>0.0002</b>	<b>0.0001</b>
		0.0292	0.0192	0.0154	0.0123	0.0006	0.0004	0.0003	0.0002
		0.0389	0.0234	0.0177	0.0132	0.0008	0.0005	0.0003	0.0003
		<b>0.0256</b>	<b>0.0200</b>	<b>0.0171</b>	<b>0.0147</b>	<b>0.0005</b>	<b>0.0004</b>	<b>0.0003</b>	<b>0.0003</b>
		0.2397	0.1542	0.1155	0.0874	0.0048	0.0031	0.0023	0.0017
	$k = 10$	0.0431	0.0287	0.0229	0.0184	0.0009	0.0006	0.0005	0.0004
		0.1079	0.0600	0.0435	0.0331	0.0021	0.0012	0.0009	0.0007
		0.2498	0.0956	0.0584	0.0390	0.0050	0.0019	0.0012	0.0008
		<b>0.0878</b>	<b>0.0619</b>	<b>0.0512</b>	<b>0.0428</b>	<b>0.0018</b>	<b>0.0012</b>	<b>0.0010</b>	<b>0.0009</b>
		0.7786	0.4504	0.3171	0.2310	0.0156	0.0090	0.0063	0.0046
	$k = 25$	0.4062	0.2003	0.1355	0.0966	0.0081	0.0040	0.0027	0.0019
		1.1483	0.5773	0.3361	0.2148	0.0230	0.0115	0.0067	0.0043
		2.6215	0.9502	0.5253	0.3290	0.0524	0.0190	0.0105	0.0066
		<b>0.2280</b>	<b>0.1200</b>	<b>0.0915</b>	<b>0.0747</b>	<b>0.0046</b>	<b>0.0024</b>	<b>0.0018</b>	<b>0.0015</b>
		1.4380	0.7473	0.4825	0.3379	0.0288	0.0149	0.0096	0.0068
	$k = 40$	2.8277	0.7556	0.4309	0.2848	0.0565	0.0151	0.0086	0.0057
		7.3905	2.0316	1.2234	0.7575	0.1478	0.0406	0.0245	0.0151
		52.933	3.3593	1.6674	0.9872	1.0587	0.0672	0.0333	0.0197

Notes: The smallest values are depicted in bold. The five estimators are the (objective) Bayesian estimator (first line in each panel), the estimator resulting from the (extended) Black–Litterman model (second line in each panel), the sample estimator (third line in each panel), the robust MVE estimator (fourth line in each panel), and the robust MCD estimator (fifth line in each panel) for several values of  $k \in \{5, 10, 25, 40\}$  and  $n \in \{50, 75, 100, 130\}$  with  $\gamma = 50$ . The asset returns are assumed to be conditionally normally distributed with large variances.

Table 3. Average absolute deviation (AD) of the estimated portfolio expected return and of the estimated portfolio variance from their population values.

$\gamma = 50$		AD of portfolio expected return				AD of portfolio variance			
Bayesian									
Black–Litterman									
Sample		$n = 50$	$n = 75$	$n = 100$	$n = 130$	$n = 50$	$n = 75$	$n = 100$	$n = 130$
Robust, MVE									
Robust, MCD									
$k = 5$		<b>0.1490</b>	<b>0.1209</b>	<b>0.1088</b>	<b>0.0956</b>	<b>0.0030</b>	<b>0.0024</b>	<b>0.0022</b>	<b>0.0019</b>
		1.2668	0.8216	0.6116	0.4632	0.0253	0.0164	0.0122	0.0093
		0.1971	0.1480	0.1268	0.1076	0.0039	0.0030	0.0025	0.0022
		0.7206	0.5815	0.5211	0.4857	0.0144	0.0116	0.0104	0.0097
		0.7974	0.6005	0.5240	0.4836	0.0159	0.0120	0.0105	0.0097
$k = 10$		<b>0.3728</b>	<b>0.3008</b>	<b>0.2633</b>	<b>0.2272</b>	<b>0.0075</b>	<b>0.0060</b>	<b>0.0053</b>	<b>0.0045</b>
		3.3112	2.0669	1.5232	1.1441	0.0662	0.0413	0.0305	0.0229
		0.7125	0.4794	0.3833	0.3089	0.0143	0.0096	0.0077	0.0062
		2.3178	1.6962	1.4892	1.3501	0.0464	0.0339	0.0298	0.0270
		3.6668	2.1039	1.6832	1.4458	0.0733	0.0421	0.0337	0.0289
$k = 25$		<b>1.4528</b>	<b>1.1011</b>	<b>0.9058</b>	<b>0.7880</b>	<b>0.0291</b>	<b>0.0220</b>	<b>0.0181</b>	<b>0.0158</b>
		11.873	6.8656	4.7531	3.4785	0.2374	0.1373	0.0951	0.0696
		6.8816	3.5138	2.3958	1.7749	0.1376	0.0703	0.0479	0.0355
		18.522	10.176	7.0974	5.6988	0.3704	0.2035	0.1419	0.1140
		29.560	14.142	10.194	8.1919	0.5912	0.2828	0.2039	0.1638
$k = 40$		<b>3.2387</b>	<b>2.3571</b>	<b>1.9424</b>	<b>1.6140</b>	<b>0.0648</b>	<b>0.0471</b>	<b>0.0388</b>	<b>0.0323</b>
		22.5595	12.4591	8.3420	5.9158	0.4512	0.2492	0.1668	0.1183
		46.3159	12.7395	7.5904	5.1566	0.9263	0.2548	0.1518	0.1031
		117.427	33.2640	20.775	14.454	2.3485	0.6653	0.4155	0.2891
		541.115	42.8432	25.721	18.805	10.822	0.8569	0.5144	0.3761

Notes: The smallest values are depicted in bold. The five estimators are the (objective) Bayesian estimator (first line in each panel), the estimator resulting from the (extended) Black–Litterman model (second line in each panel), the sample estimator (third line in each panel), the robust MVE estimator (fourth line in each panel), and the robust MCD estimator (fifth line in each panel) for several values of  $k \in \{5, 10, 25, 40\}$  and  $n \in \{50, 75, 100, 130\}$  with  $\gamma = 50$ . The asset returns are assumed to be conditionally multivariate  $t$ -distributed with 5 degrees of freedom and small variances.

Table 4. Average absolute deviation (AD) of the estimated portfolio expected return and of the estimated portfolio variance from their population values.

$\gamma = 50$		AD of portfolio expected return				AD of portfolio variance			
Bayesian									
Black–Litterman									
Sample		$n = 50$	$n = 75$	$n = 100$	$n = 130$	$n = 50$	$n = 75$	$n = 100$	$n = 130$
Robust, MVE									
Robust, MCD									
$k = 5$		<b>0.0150</b>	<b>0.0123</b>	<b>0.0110</b>	<b>0.0096</b>	<b>0.0003</b>	<b>0.0002</b>	<b>0.0002</b>	<b>0.0002</b>
		0.1085	0.0710	0.0533	0.0405	0.0021	0.0014	0.0010	0.0008
		0.0196	0.0149	0.0127	0.0107	0.0004	0.0003	0.0003	<b>0.0002</b>
		0.0671	0.0528	0.0472	0.0436	0.0013	0.0010	0.0009	0.0009
		0.0737	0.0544	0.0475	0.0435	0.0014	0.0011	0.0009	0.0008
$k = 10$		<b>0.0374</b>	<b>0.0300</b>	<b>0.0263</b>	<b>0.0225</b>	<b>0.0007</b>	<b>0.0006</b>	<b>0.0005</b>	<b>0.0004</b>
		0.2958	0.1863	0.1386	0.1043	0.0059	0.0037	0.0028	0.0021
		0.0706	0.0474	0.0379	0.0303	0.0014	0.0009	0.0008	0.0006
		0.2242	0.1613	0.1412	0.1271	0.0045	0.0032	0.0028	0.0025
		0.3534	0.1998	0.1592	0.1361	0.0070	0.0040	0.0032	0.0027
$k = 25$		<b>0.1476</b>	<b>0.1116</b>	<b>0.0912</b>	<b>0.0792</b>	<b>0.0023</b>	<b>0.0022</b>	<b>0.0018</b>	<b>0.0015</b>
		1.0651	0.6335	0.4441	0.3277	0.0213	0.0127	0.0089	0.0065
		0.6878	0.3502	0.2383	0.1763	0.0137	0.0070	0.0048	0.0035
		1.8374	1.0074	0.6999	0.5614	0.0367	0.0201	0.0140	0.0112
		2.9172	1.3960	1.0026	0.8062	0.0583	0.0279	0.0200	0.0161
$k = 40$		<b>0.3257</b>	<b>0.2392</b>	<b>0.1970</b>	<b>0.1630</b>	<b>0.0065</b>	<b>0.0048</b>	<b>0.0039</b>	<b>0.0033</b>
		1.9528	1.1312	0.7753	0.5575	0.0391	0.0226	0.0155	0.0111
		4.6225	1.2720	0.7580	0.5142	0.0924	0.0254	0.0152	0.0103
		11.768	3.3082	2.0689	1.4391	0.2354	0.0662	0.0414	0.0288
		49.947	4.2723	2.5567	1.8670	0.9989	0.0854	0.0511	0.0373

Notes: The smallest values are depicted in bold. The five estimators are the (objective) Bayesian estimator (first line in each panel), the estimator resulting from the (extended) Black–Litterman model (second line in each panel), the sample estimator (third line in each panel), the robust MVE estimator (fourth line in each panel), and the robust MCD estimator (fifth line in each panel) for several values of  $k \in \{5, 10, 25, 40\}$  and  $n \in \{50, 75, 100, 130\}$  with  $\gamma = 50$ . The asset returns are assumed to be conditionally multivariate  $t$ -distributed with 5 degrees of freedom and large variances.



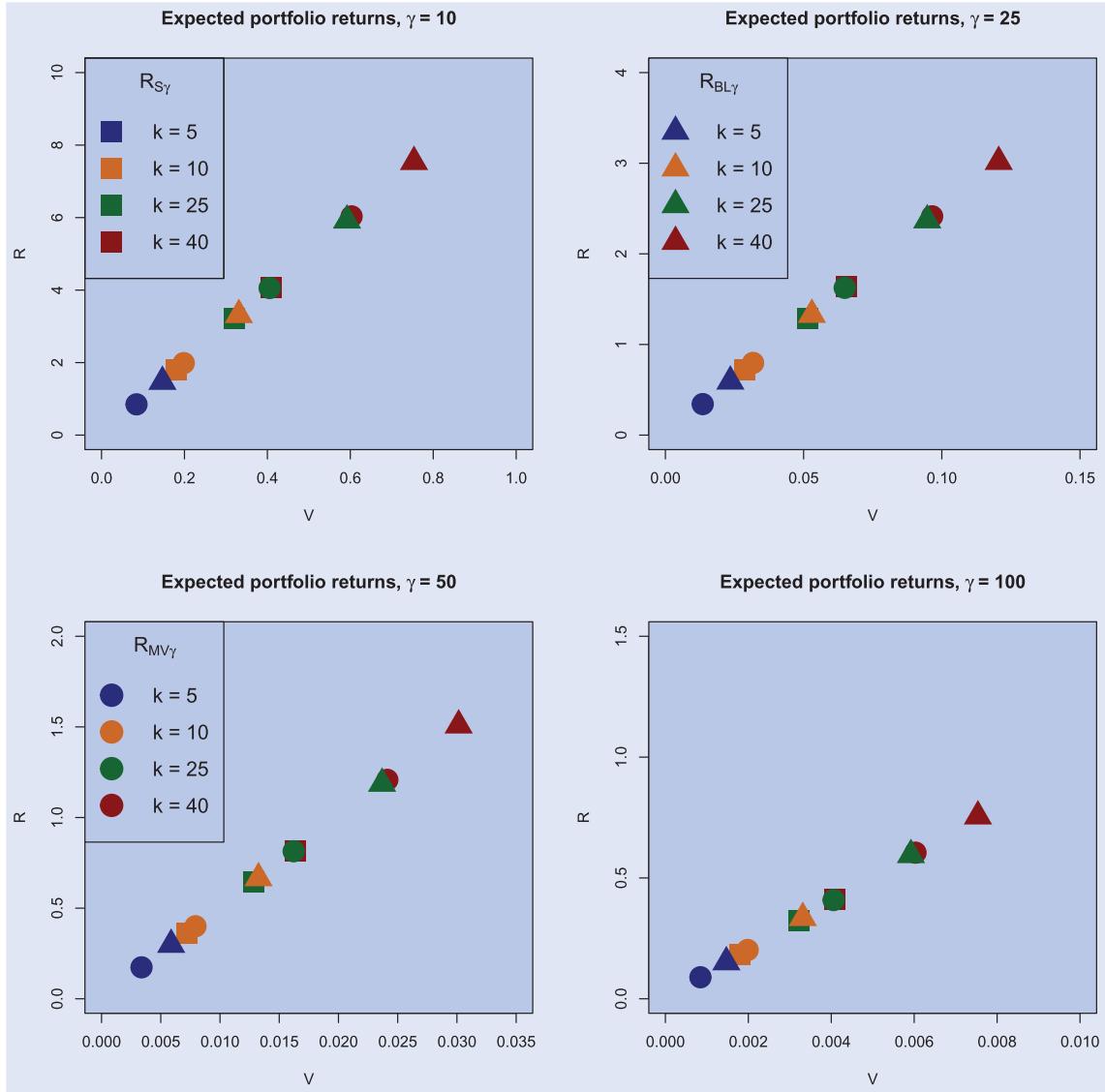


Figure 2. Sample optimal portfolios (squares), (objective) Bayesian optimal portfolios (circles), and the Black–Litterman optimal portfolios (triangles) for the risk aversion coefficient of  $\gamma \in \{10, 25, 50, 100\}$ , for the sample case of  $n = 130$  and for the portfolio dimension of  $k \in \{5, 10, 25, 40\}$  in the case of weekly data.

(1) The optimization problem

$$\mathbf{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w} \longrightarrow \max \quad \text{subject to } \mathbf{w}^\top \mathbf{1} = 1 \quad (27)$$

is solved resulting in the expression of optimal portfolio weights presented in terms of the population (unknown) parameters  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$ :

$$\begin{aligned} \mathbf{w}_{P,\gamma} &= \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}} + \gamma^{-1} \mathbf{R} \boldsymbol{\mu} \quad \text{with} \\ \mathbf{R} &= \boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1} \mathbf{1}' \boldsymbol{\Sigma}^{-1}}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}} \end{aligned} \quad (28)$$

with the expected return and the variance expressed as

$$\begin{aligned} R_{P,\gamma} &= \frac{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}} + \gamma^{-1} \boldsymbol{\mu}^\top \mathbf{R} \boldsymbol{\mu} \quad \text{and} \\ V_{P,\gamma} &= \frac{1}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}} + \gamma^{-2} \boldsymbol{\mu}^\top \mathbf{R} \boldsymbol{\mu}. \end{aligned} \quad (29)$$

(2) The unknown population quantities are replaced by their sample counterparts, i.e. by the sample mean vector and the sample covariance matrix given by

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}_{t-1} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = d_n \mathbf{S}_{t-1} \quad \text{with} \quad d_n = \frac{1}{n-1}.$$

Then the sample optimal portfolio weights are obtained by

$$\mathbf{w}_{S,\gamma} = \frac{\mathbf{S}_{t-1}^{-1} \mathbf{1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} + \gamma^{-1} d_n^{-1} \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1} \quad (30)$$

with the sample estimators for the expected return and for the variance given by

$$\begin{aligned} R_{S,\gamma} &= \frac{\mathbf{1}^\top \mathbf{S}_{t-1}^{-1} \bar{\mathbf{x}}_{t-1}}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} + \gamma^{-1} d_n^{-1} \bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1} \quad \text{and} \\ V_{S,\gamma} &= \frac{d_n}{\mathbf{1}' \mathbf{S}_{t-1}^{-1} \mathbf{1}} + \gamma^{-2} d_n^{-1} \bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1}. \end{aligned} \quad (31)$$

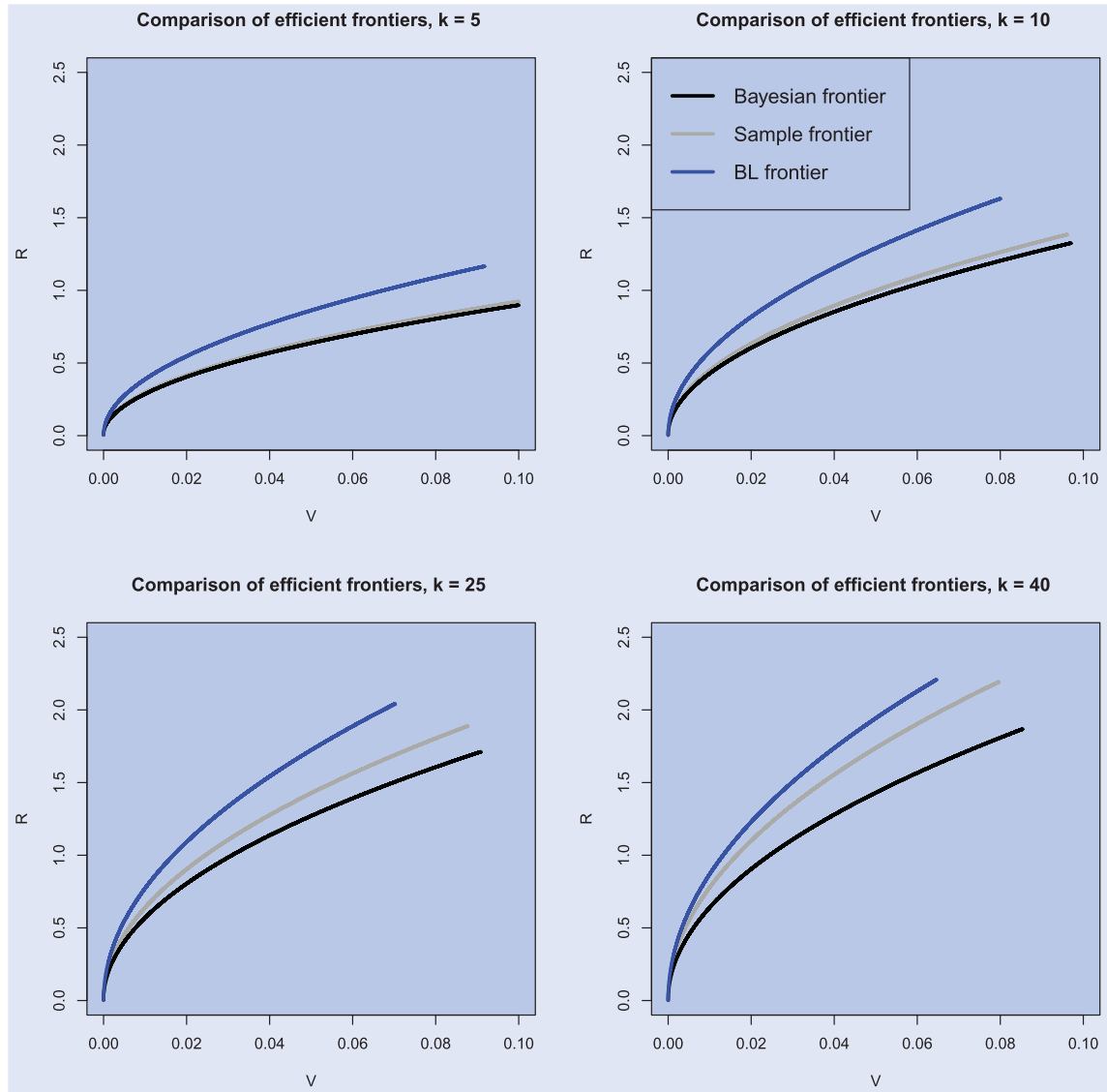


Figure 3. The sample efficient frontier, the (objective) Bayesian efficient frontier, and the Black–Litterman efficient frontier for  $n = 130$  and  $k \in \{5, 10, 25, 40\}$  in the case of weekly data.

In the similar way, the sample efficient frontier is constructed by (see Bodnar and Schmid 2008, 2009, Kan and Smith 2008) and it is expressed as

$$(R - R_{GMV,S})^2 = \frac{\bar{\mathbf{x}}_{t-1}^\top \mathbf{Q}_{t-1} \bar{\mathbf{x}}_{t-1}}{d_n} (V - V_{GMV,S}), \quad (32)$$

where

$$R_{GMV,S} = \frac{\mathbf{1}^\top \mathbf{S}_{t-1}^{-1} \bar{\mathbf{x}}_{t-1}}{\mathbf{1}^\top \mathbf{S}_{t-1}^{-1} \mathbf{1}} \quad \text{and} \quad V_{GMV,S} = \frac{d_n}{\mathbf{1}^\top \mathbf{S}_{t-1}^{-1} \mathbf{1}}. \quad (33)$$

Formula (32) presents the sample estimator of the population efficient frontier.

### 3.2. Comparison of the three estimators of the efficient frontier

It is remarkable that the expression of the sample optimal portfolio weights has the same structure as the weights of the optimal portfolios obtained following the (objective) Bayesian

approach. The only difference is that  $c_{k,n}$  in (8) is replaced by  $d_n$  in (30). Similar results are also obtained in the case of the efficient frontier which is fully determined by three parameters: the mean and the variance of the global minimum variance portfolio and the slope parameter. While the formulae in the case of the mean of the global minimum variance portfolio coincide, this is not longer true for the variance of the global minimum variance portfolio and the slope coefficient. The (objective) Bayesian approach leads to a larger value of the variance and to a smaller value of the slope parameter. The difference between the corresponding expressions obtained by the sample estimation or derived from the Bayesian posterior distribution as in Section 2 can be considerable when the portfolio dimension is comparable to the sample size as shown in figure 1, where we plot the ratio  $c_{k,n}/d_n$  as a function of  $k/n$  for  $n \in \{50, 100\}$ . We observe that when the number of assets  $k$  gets closer to the sample size, even for a moderate ratio of  $k/n = 0.6$ , the (objective) Bayesian estimator and the sample estimator deviate. If the number of assets corresponds almost to the sample size, the

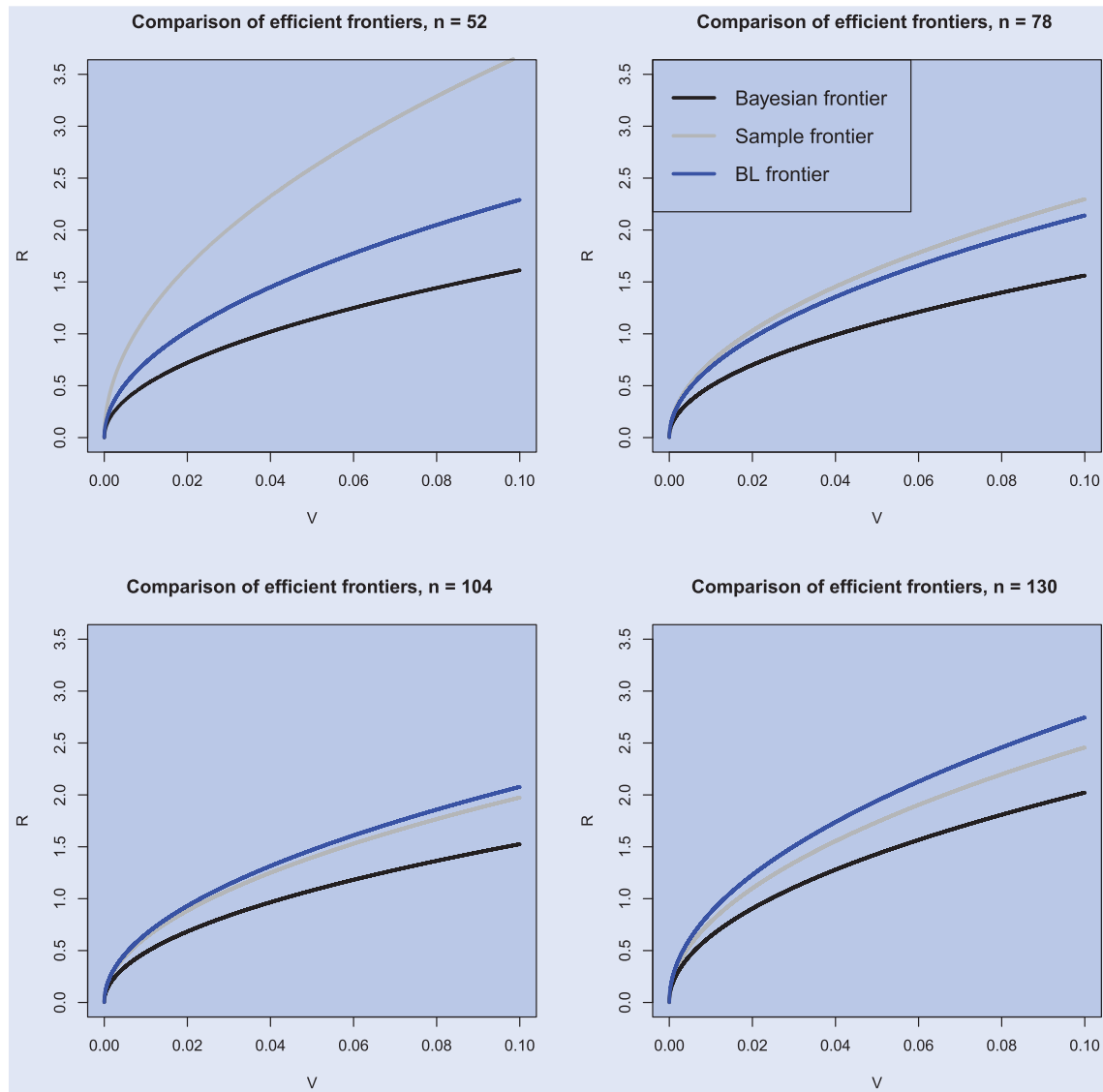


Figure 4. The sample efficient frontier, the (objective) Bayesian efficient frontier, and the Black–Litterman efficient frontier for  $k = 40$  and  $n \in \{52, 78, 104, 130\}$  in the case of weekly data.

estimators deviate considerably. Since it is sometimes necessary to restrict an estimation to a smaller sample size, e.g. after a structural break in the data, the difference in the estimators has to be considered. To this end, we note that such a simple comparison of the estimated efficient frontiers cannot be performed in the case of the Black–Litterman efficient frontier due to a more complicated structure of the latter which also depends on the expert knowledge about the parameters of the asset return distribution.

It is a well-known fact that the sample efficient frontier is overoptimistic and overestimates the location of the population efficient frontier in the mean–variance space (cf. Basak *et al.* 2005, Siegel and Woodgate 2007, Bodnar and Bodnar 2010). In contrast, the Bayesian approach provides an improved procedure which shrinks the sample efficient frontier by increasing the estimated variance of the global minimum portfolio and reducing the slope parameter. We will illustrate this point in Section 4 on real data described in Section 4.1.

### 3.3. Simulation study

We provide a detailed comparison of the estimators of the optimal portfolio weights, namely the suggested (objective) Bayesian approach, the estimator resulting from the (extended) Black–Litterman model, and the sample estimator, via simulations in this section. In the comparison study we also include two robust estimators of optimal portfolio weights (see, Chapter 20 in Würtz *et al.* 2015), which are based on the robust estimation of the mean vector and of the covariance matrix known as the minimum volume ellipsoid (MVE) estimator (see, e.g. Rousseeuw 1984) and the minimum covariance determinant (MCD) estimator (see, e.g. Rousseeuw and Driessen 1999). The aim of the Monte Carlo study is to assess the performance of each strategy in the estimation of the expected return and the variance of optimal portfolios. Such results will provide a better understanding about potential improvements which can be obtained by employing the new Bayesian approach.

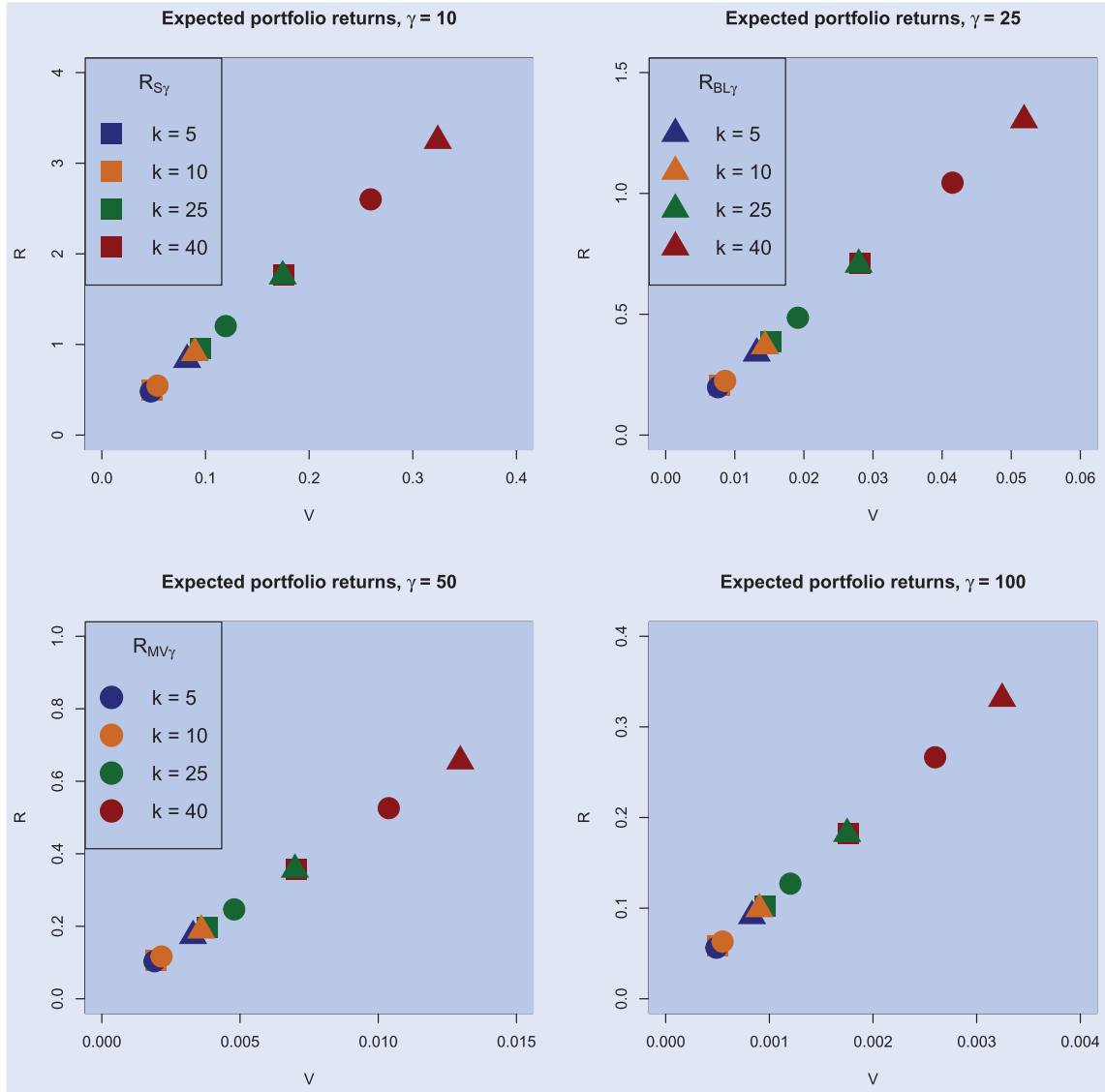


Figure 5. Sample optimal portfolios (squares), (objective) Bayesian optimal portfolios (circles), and the Black–Litterman optimal portfolios (triangles) for the risk aversion coefficient of  $\gamma \in \{10, 25, 50, 100\}$ , for the sample case of  $n = 130$  and for the portfolio dimension of  $k \in \{5, 10, 25, 40\}$  in the case of monthly data.

The results of Proposition 4.6 of Bernardo and Smith (2000) ensure that the conditional multivariate normal distribution satisfies the imposed assumptions of infinitely exchangeability and of multivariate centred spherically symmetry. For that reason, we assume that the asset returns are independently and identically distributed given mean vector  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)^\top$  and covariance matrix  $\boldsymbol{\Sigma}$  with the conditional distribution given by  $\mathbf{X}_t | \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . In order to avoid any restriction to specific values of  $\boldsymbol{\mu}$ , the elements of this vector were generated from the uniform distribution on  $[-0.01, 0.01]$  in each simulation run, that is  $\mu_i \sim \text{Unif}(-0.01, 0.01)$ . For the covariance matrix we consider its decomposition into the correlation matrix  $\mathbf{R}$  and the diagonal matrix with standard deviations  $\mathbf{D} = \text{diag}(d_1, \dots, d_k)$ , i.e.  $\boldsymbol{\Sigma} = \mathbf{R}\mathbf{D}\mathbf{D}\mathbf{R}$ . Two choices of volatility are considered: (i) low volatility with  $d_i \sim \text{Unif}(0.002, 0.005)$  and (ii) high volatility with  $d_i \sim \text{Unif}(0.005, 0.02)$ . The correlation matrix is set to  $\mathbf{R} = (1 - \rho)\mathbf{I}_k + \rho\mathbf{J}_k$  with  $\rho = 0.6$ ,  $k$ -dimensional identity matrix  $\mathbf{I}_k$ , and the  $k$ -dimensional matrix of ones  $\mathbf{J}_k$ . We put

$k \in \{5, 10, 25, 40\}$ ,  $n \in \{50, 75, 100, 130\}$ , and  $\gamma = 50$ . In the case of the (extended) Black–Litterman model the precision parameters are  $r_0 = 100$  and  $d_0 = 100$ , while  $\mathbf{m}_0$  and  $\mathbf{S}_0$  are obtained by perturbing  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  as  $\mathbf{m}_0 = \boldsymbol{\mu} + 0.5\boldsymbol{\varepsilon}$  and  $\mathbf{S}_0 = \boldsymbol{\Sigma} + 0.5\boldsymbol{\Delta}$  with  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_k)^\top$  and  $\boldsymbol{\Delta} = \text{diag}(\delta_1^2, \dots, \delta_k^2)$  where  $\varepsilon_i \sim \text{Unif}(-0.01, 0.01)$  and  $\delta_i \sim \text{Unif}(0.001, 0.005)$ . The results in the tables are based on  $B = 10\,000$  independent repetitions.

As a measure of performance, the average absolute deviation from the resulting estimator to the corresponding true population value was computed for the portfolio expected return and the portfolio variance. The values are summarized in table 1 in the case of low volatilities and in table 2 for high volatilities. We observe that the application of the new objective Bayesian estimation strategy leads to the considerable improvements in terms of both performance measures meaning a better point estimation of both the portfolio expected return and the portfolio variance. The impact of the improvement increases as the portfolio dimension becomes larger.

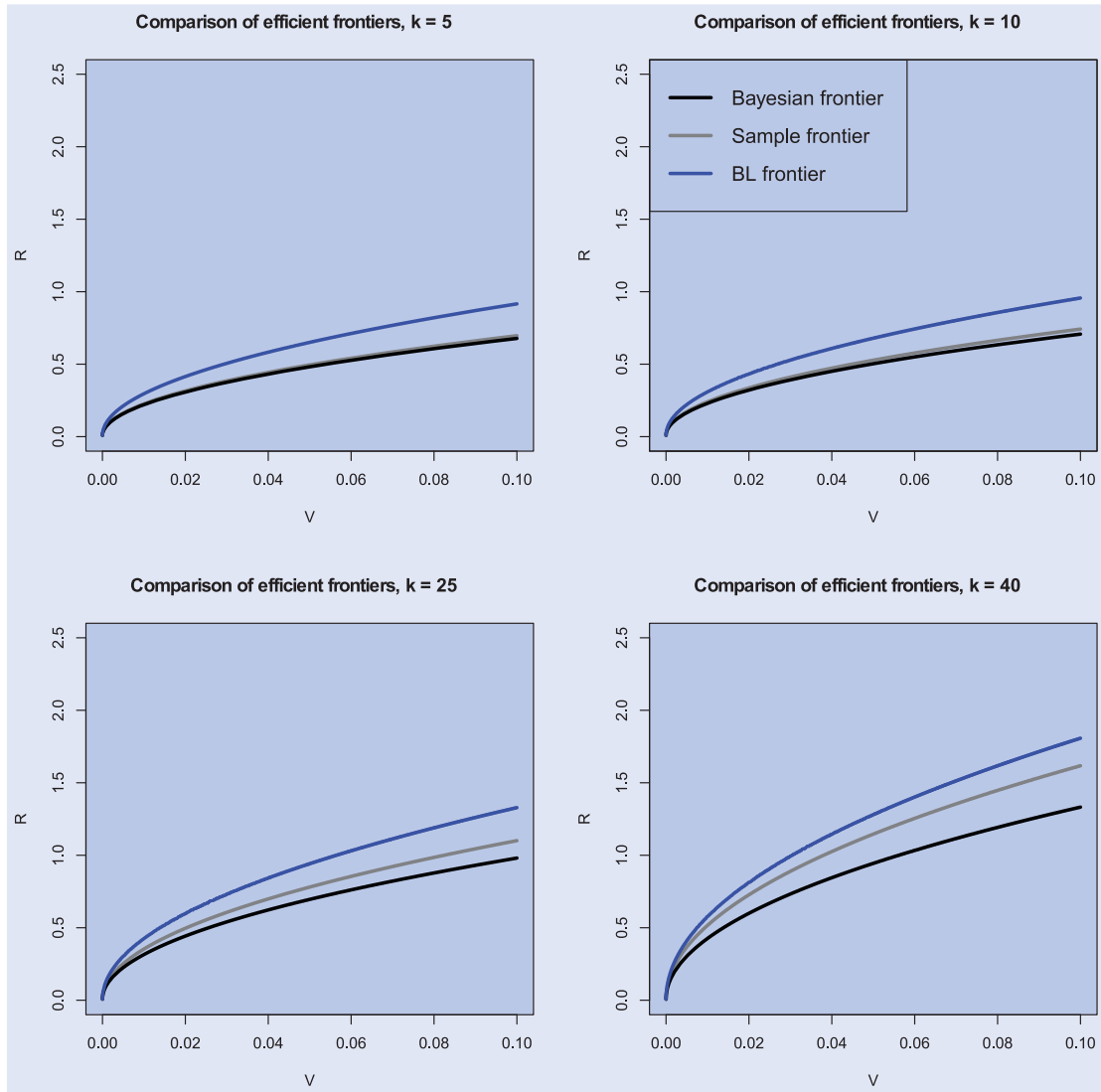


Figure 6. The sample efficient frontier, the (objective) Bayesian efficient frontier, and the Black–Litterman efficient frontier for  $n = 130$  and  $k \in \{5, 10, 25, 40\}$  in the case of monthly data.

Especially, when  $k = 40$  and  $n = 50$  the new Bayesian estimator results in the values of the average deviation which are 12 times smaller than the one computed for the sample portfolio in the case of the portfolio expected return and about 11.7 times smaller in the case of the variance when the volatilities are low, while both these values are above 12.2 for high volatilities. These findings are in line with the results presented in figure 1. Also, a slightly better performance is observed in the case of the (extended) Black–Litterman estimation strategy when the portfolio dimension is large. In this case, it is ranked on the second place by using both criteria when  $k = 40$  and  $n = 50$ , while it is on the third place in all other cases showing that the influence of the expert knowledge could have a great impact when the sample size is not large. Similar findings are also present for two robust portfolio selection strategies, which are ranked on the fourth and on the fifth places. Also, in these cases, the sample size is not large enough with respect to the portfolio dimension, which causes a bad performance of these two strategies. Finally, we point out that with the increase of the sample size, the values of the two performance measures becomes smaller and in

the case of  $k = 5$  and  $n = 130$  they are almost the same for the new Bayesian approach and the sample method while the Bayesian approach is still more preferable.

### 3.4. Robustness analysis

Next, we investigate how robust are the numerical findings obtained in the previous section to the deviation from the conditional normality. For this purposed, we employ the conditional multivariate  $t$ -distribution with 5 degrees of freedom, which has the same mean vector  $\mu$  and the same covariance matrices  $\Sigma$  as in the case of the model from Section 3.3. Furthermore, it is noted that in contrast to the conditional multivariate normal distribution, the conditional multivariate  $t$ -distribution does not belong to the family of infinitely exchangeability and multivariate centred spherically symmetrical distributions.

The replacement of the conditional multivariate normal distribution by the conditional multivariate  $t$ -distribution influences the values of the average absolute deviation computed in both cases of the portfolio expected return and of



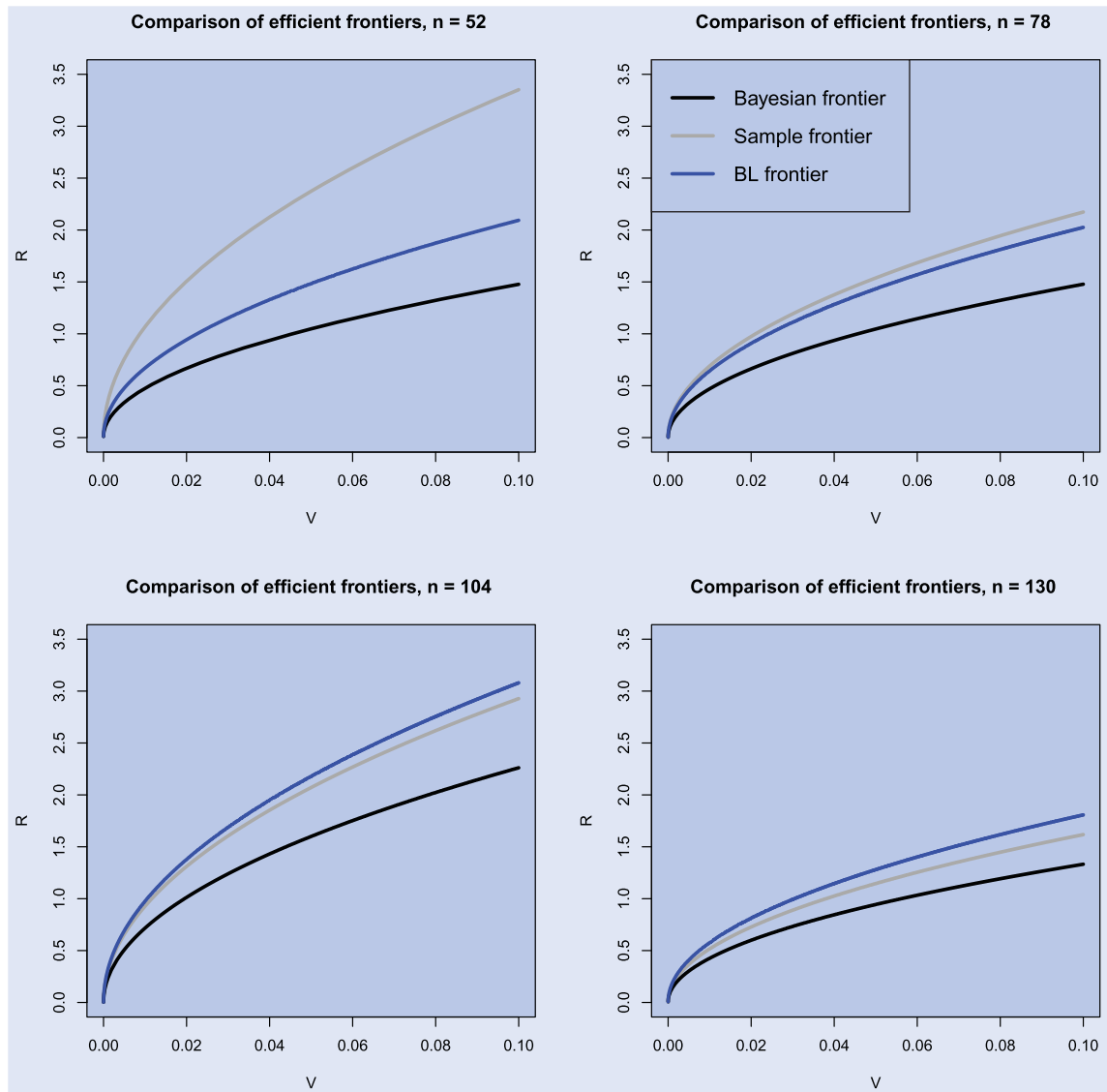


Figure 7. The sample efficient frontier, the (objective) Bayesian efficient frontier, and the Black–Litterman efficient frontier for  $k = 40$  and  $n \in \{52, 78, 104, 130\}$  in the case of monthly data.

the portfolio variance. All these values become considerably larger which is explained by the heavy-tailed nature of the multivariate  $t$ -distribution (see tables 3 and 4). On the other side, the ranking between the five estimation strategies does not change. The new (objective) Bayesian approach outperforms the other four competitors in all of the considered cases similarly when the observation data were generated from the multivariate normal distribution. Also, in the case of the large-dimensional portfolio consisting of 40 assets and the sample size equal to  $n = 50$ , the (extended) Black–Litterman approach is ranked on the second place for both low and high volatilities, while the sample estimator performs better in the rest of the considered cases.

## 4. Empirical illustration

### 4.1. Data

For the first empirical illustration, we use weekly returns from a collection of assets of the S&P500, allowing for portfolios

ranging from 5 to 40 assets. A similar setup is also used in the second empirical illustration where monthly returns instead of weekly returns are used. The parameters are estimated with sample sizes of  $n \in \{52, 78, 104, 130\}$ , corresponding to one year up to two and a half years of weekly data or to approximately four and a half up to eleven years of monthly data. All the data end on the 8th of October 2017. The constructed portfolios consist of  $k \in \{5, 10, 25, 40\}$  assets. The hyperparameters in the extended Black–Litterman model are obtained by employing the empirical Bayes approach (see, e.g. Gelman *et al.* 2014, Bauder *et al.* 2020). This allows us to analyse the behaviour of the proposed model not only in terms of economic risk but also regarding statistical estimation uncertainty.

### 4.2. Results for weekly data

As mentioned in Section 3, there is a distinct difference between the classical sample estimators and the (objective)

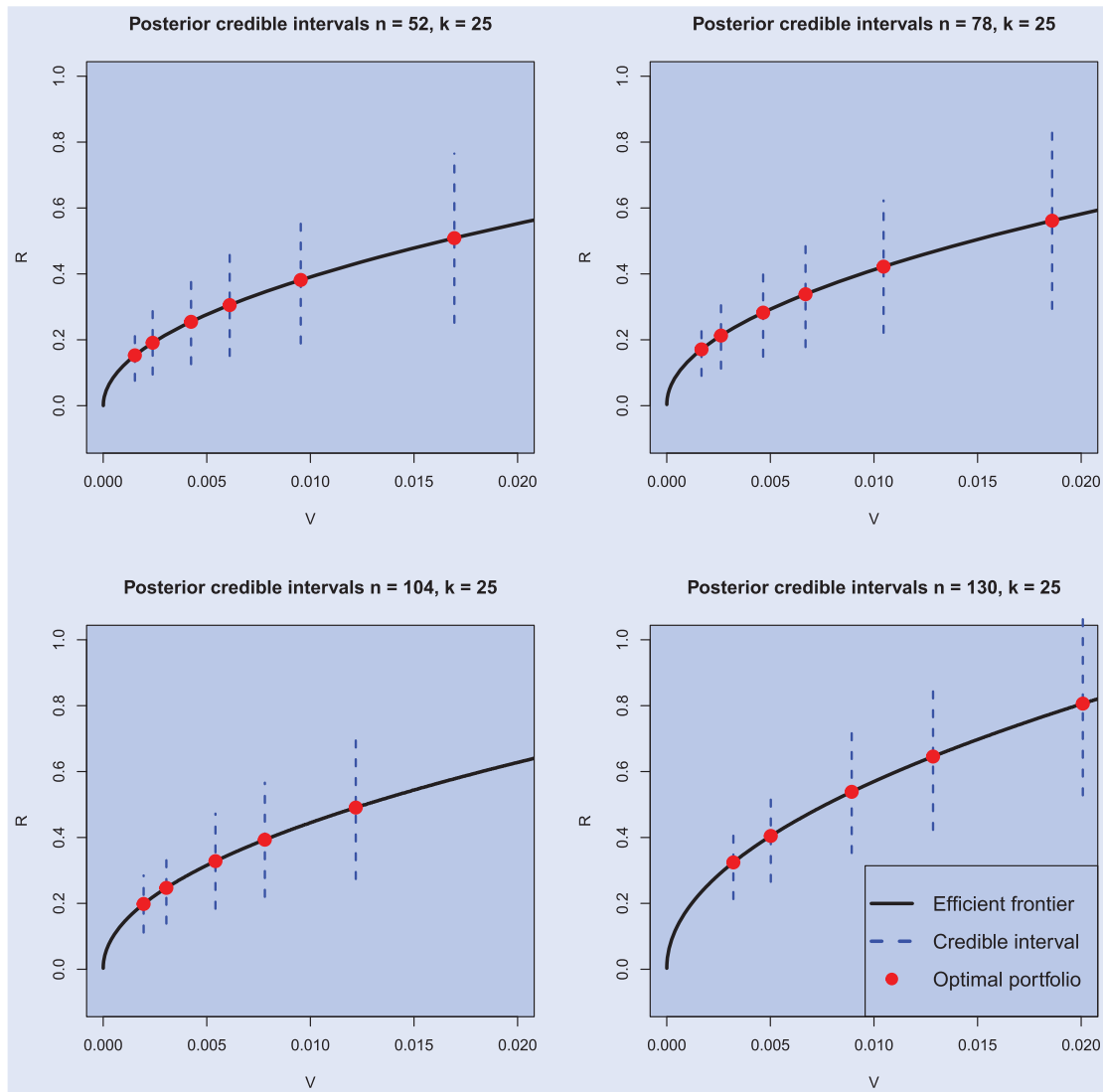


Figure 8. Credible intervals for the return of optimal portfolios with varying risk attitudes for weekly data obtained by employing the (objective) Bayesian approach. The sample sizes are chosen to be  $n \in \{52, 78, 104, 130\}$  and the portfolio dimension is fixed to  $k = 25$ . The confidence level is set to  $\alpha = 0.05$ .

Bayesian estimators proposed in this paper. With this conclusion and the fact that the sample efficient frontier overestimates the population efficient frontier, we expect the estimators for the expected return and the variance to be larger in the Bayesian case compared to the sample estimators indicating that the (objective) Bayesian approach also takes the estimation risk into account in its construction which in practice automatically leads to smaller values of the risk aversion coefficient in comparison to the conventional case. Figure 2 illustrates this presumption: fixing  $n = 130$  and considering different portfolio sizes  $k \in \{5, 10, 25, 40\}$  for different risk attitudes  $\gamma \in \{10, 25, 50, 100\}$ , we find that for the same value of the risk coefficient  $\gamma$  and for the same portfolio size, the (objective) Bayesian estimator performs as expected compared to the sample estimator, whereas the Black–Litterman optimal portfolios exhibit a more exaggerated behaviour. The latter results are related to the usage of the additional information in the construction of the Black–Litterman optimal portfolios which, in particular, can lead to

the increase of uncertainty especially when the hyperparameters differ considerably from the corresponding population values as shown in the simulation study of Section 3. Furthermore, the difference in the estimators increases if the number of assets gets closer to the sample size, as illustrated in figure 1 or when  $\gamma$  decreases, i.e. for less risk averse investors the impact of parameter uncertainty becomes larger.

Regarding the efficient frontier, figure 3 shows the estimated efficient frontiers for a fixed sample size of  $n = 130$  and varying portfolio sizes  $k \in \{5, 10, 25, 40\}$  in all three cases, namely the sample efficient frontier, the (objective) Bayesian efficient frontier, and the Black–Litterman efficient frontier. The (objective) Bayesian efficient frontier lies always below the sample efficient frontier and therefore exhibits less overestimation of the population efficient frontier. In contrast, the Black–Litterman frontier exhibits even a stronger overestimation compared to the population efficient frontier due to the uncertainty related to hyperparameters which are present in the model. Furthermore, figure 3 also illustrates the finding

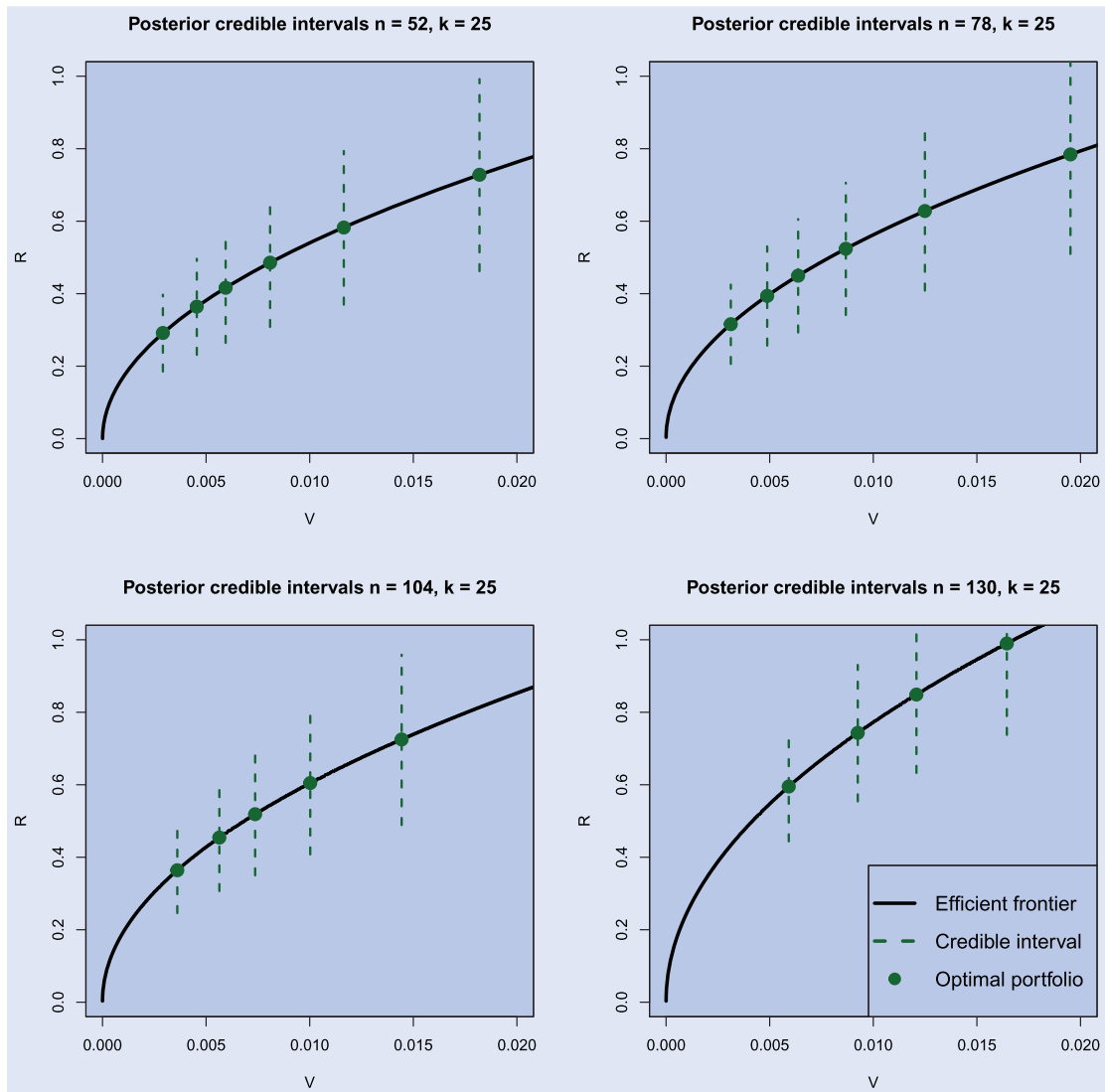


Figure 9. Credible intervals for the return of optimal portfolios for the Black–Litterman model with varying risk attitudes for weekly data. The sample sizes are chosen to be  $n \in \{52, 78, 104, 130\}$  and the portfolio dimension is fixed to  $k = 25$ . The confidence level is set to  $\alpha = 0.05$ .

shown in figure 1. The estimators of the efficient frontier deviate stronger when the portfolio size gets closer to the sample size. This fact is also illustrated in figure 4 for fixed  $k = 40$  and varying  $n \in \{52, 78, 104, 130\}$ . The (objective) Bayesian and the sample estimated efficient frontiers coincide more the larger the sample size  $n$  is, whereas the Black–Litterman efficient frontier appears to exhibit stronger overestimation with growing sample size. This is in line with the theoretical implications. Finally, we also observe the increase in the slope parameter of the efficient frontier when the portfolio dimension increases indicating the well-documented positive effect of portfolio diversification.

#### 4.3. Results for monthly data

Figure 5 shows the location of the sample optimal portfolios, of the (objective) Bayesian optimal portfolios, and of the Black–Litterman optimal portfolios computed for the same values of the risk aversion coefficient  $\gamma$ , portfolio

dimension  $k$ , and sample size  $n$  as in figure 2 in the case of monthly data. The distinct difference between the classical sample estimators, the (objective) Bayesian estimators, and the Black–Litterman optimal portfolios is also identified for monthly data. In contrast to figure 2 we observe a considerable reduction in both the expected returns and the variances of all constructed optimal portfolios, while their ordering with respect to the location in the mean–variance space is the same as the one observed in figure 2. For the same value of the investor risk aversion coefficient the sample optimal portfolios exhibit smaller values of the expected return and the variance following by the (objective) Bayesian optimal portfolio which incorporate the parameter uncertainty into account in their construction. Finally, the uncertainty about the hyperparameters move the Black–Litterman optimal portfolios further in the direction of larger values of the expected return and variance.

Similar findings are also present in figures 6 and 7 where the sample efficient frontier, the (objective) Bayesian efficient

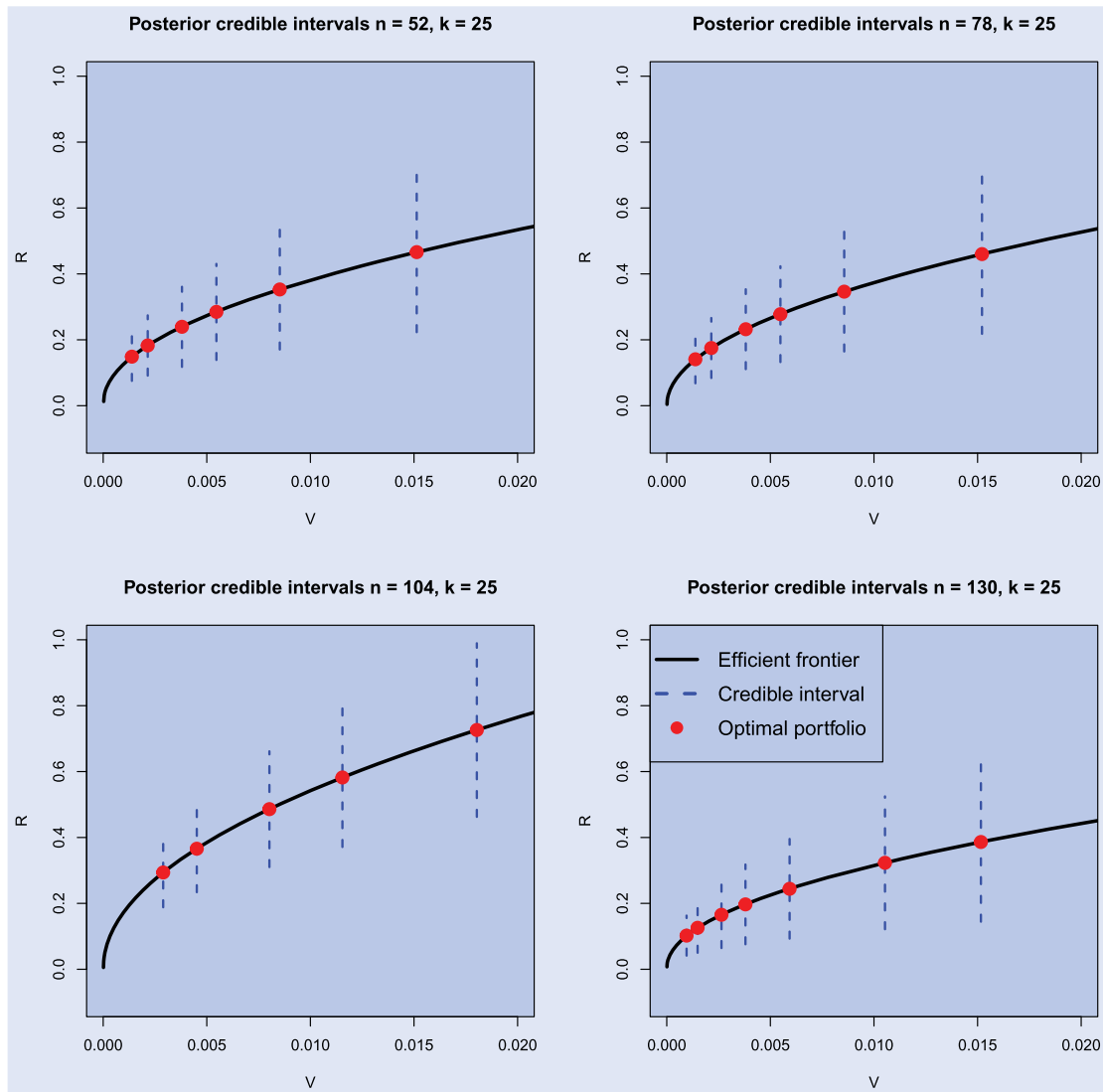


Figure 10. Credible intervals for the return of optimal portfolios with varying risk attitudes for monthly data obtained by employing the (objective) Bayesian approach. The sample sizes are chosen to be  $n \in \{52, 78, 104, 130\}$  and the portfolio dimension is fixed to  $k = 25$ . The confidence level is set to  $\alpha = 0.05$ .

frontier, and the Black–Litterman efficient frontier are drawn for several values of portfolio dimension  $k$  and sample size  $n$ . Both the sample efficient frontier and the Black–Litterman efficient frontier lie above the (objective) Bayesian efficient frontier tending to provide a considerable overestimation of the population efficient frontier especially when the portfolio dimension is large in comparison to the sample size. In figure 7 we also observe that sample efficient frontier is located above the other two estimators when the sample size is only slightly larger than portfolio dimension  $k = 40$  indicating its poor performance in such situations independently of the data frequency used in the estimation. On the other side, the Black–Litterman efficient frontier demonstrates its dependence on the chosen hyperparameters used in its construction. The considerable sample sizes in both figures 6 and 7 seem to be not large enough to reduce the effect of the hyperparameters on the resulting estimator of the efficient frontier. Better results are expected for larger sample sizes following the Bernstein–von Mises theorem (cf. Bernardo and Smith 2000).

#### 4.4. Posterior interval prediction

In contrast to the conventional procedure, both the (objective) Bayesian approach and the application of the Black–Litterman model provide also the whole posterior predictive distribution of the optimal portfolio return and not only the point estimator of its weights. Using data described in Section 4.1, we calculate in this section the prediction intervals for the optimal portfolio returns for several values of the risk-aversion coefficient  $\gamma \in \{10, 20, \dots, 100\}$ , for  $k \in \{5, 25\}$ , and for  $n \in \{52, 78, 104, 130\}$  in the case of weekly data (see, figures 8 for the (objective) Bayesian approach and 9 for the Black–Litterman model) and in the case of monthly data (see, figures 10 for the (objective) Bayesian approach and figure 11 for the Black–Litterman model).

The prediction intervals in figures 8 – 11 are obtained by using the following procedure:

- (a) Fix  $\gamma$  and calculate the expected return and the variance of the corresponding mean–variance optimal portfolio as given (9) and (10) for the (objective)

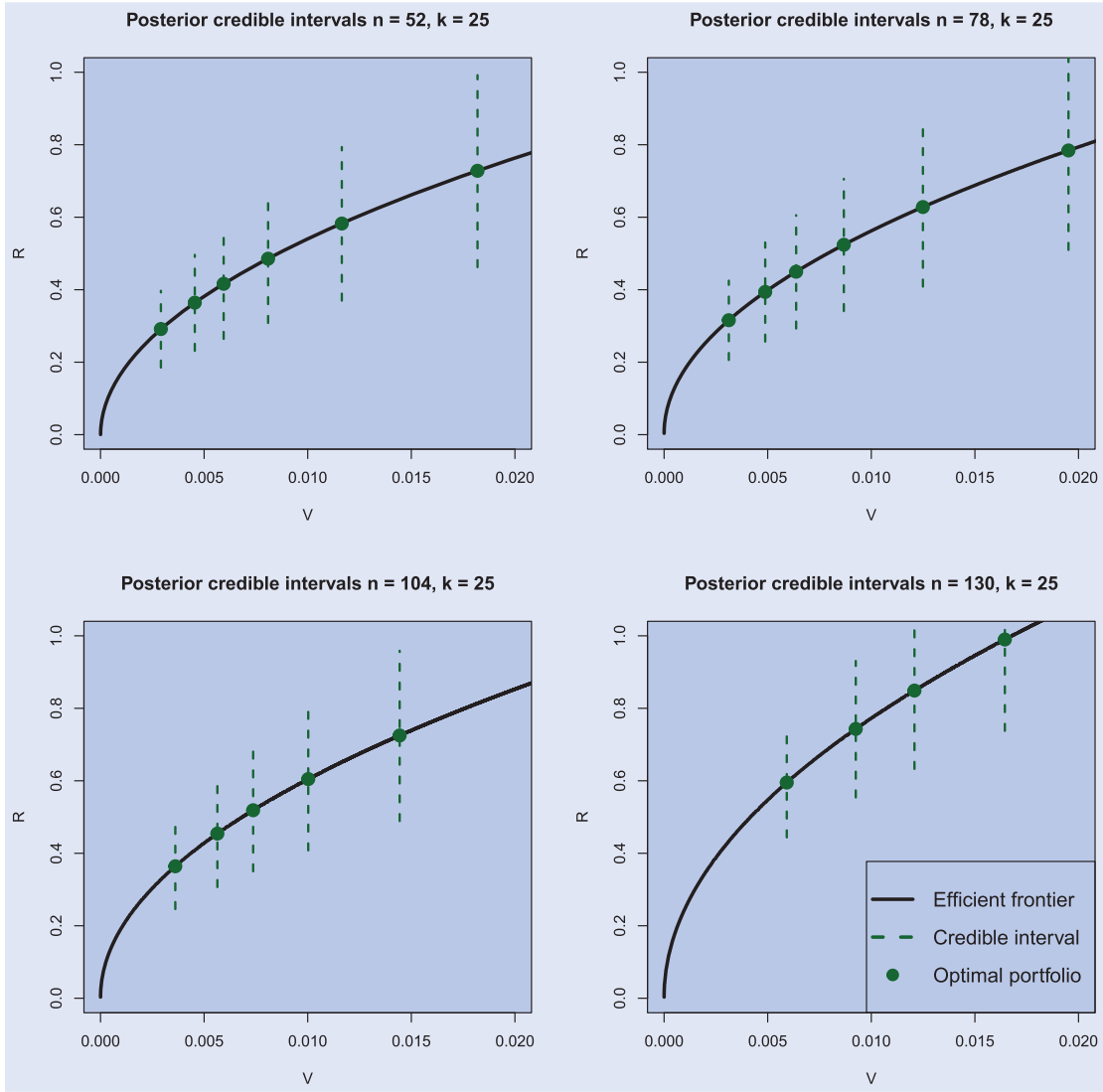


Figure 11. Credible intervals for the return of optimal portfolios for the Black–Litterman model with varying risk attitudes for monthly data. The sample sizes are chosen to be  $n \in \{52, 78, 104, 130\}$  and the portfolio dimension is fixed to  $k = 25$ . The confidence level is set to  $\alpha = 0.05$ .

- Bayesian approach and in (22) and (23) for the Black–Litterman model;
- (b) For chosen  $\gamma$ , compute the weights of the optimal mean–variance portfolio  $\mathbf{w}_{MV,\gamma}$  using (8) for the (objective) Bayesian approach and (21) for the Black–Litterman model;
- (c) In using  $\mathbf{w}_{MV,\gamma}$  apply the results of Theorem 1 and Algorithm 1 for the (objective) Bayesian approach and Theorem 2 and Algorithm 2 for the Black–Litterman model to get a sample of optimal portfolio returns denoted by  $R_{MV,\gamma}^{(b)}$  for  $b = 1, \dots, B$ ;
- (d) Fix the significance level of the prediction interval  $\alpha$  and compute the  $\alpha/2$ - and  $(1 - \alpha/2)$ -quantiles from the empirical distribution of  $R_{MV,\gamma}^{(b)}$ ,  $b = 1, \dots, B$  in both cases of the (objective) Bayesian approach and of the Black–Litterman model;
- (e) For the computed value of  $V_{MV,\gamma}$  in part (a), plot the point prediction  $R_{MV,\gamma}$  from (a) together with the prediction interval from (d). This procedure is performed separately for the sample  $R_{MV,\gamma}^{(b)}$  for

$b = 1, \dots, B$  obtained by employing the (objective) Bayesian approach and for a similar sample obtained by using the Black–Litterman model.

The order of the efficient portfolios given in figures 8–11 is directly determined by the risk aversion coefficient. The smaller  $\gamma$ , the riskier is the portfolio and lies therefore more right on the efficient frontier. We observe that the optimal efficient portfolios are shifted to the right for growing sample sizes. But the focus lies here on the credible intervals for a confidence level of  $\alpha = 0.05$ . The first observation is that no credible interval covers negative values, implying positive portfolio returns with probability of 95%. The second observation is that the credible intervals become larger the more risky an efficient portfolio becomes—which is in line with the theory. And the third observation is that these credible intervals for riskier efficient portfolios become larger regardless of the increased sample size. Hence, the decrease in estimation risk resulting from a larger sample is outweighed by the economic risk.



As in the previous section, no differences between the results obtained by using weekly data or monthly data are detected. Most noteworthy is the case of  $n = 130$  for monthly data as shown in figures 10 and 11, covering the financial crisis of 2008, showing a drastic drop in returns compared to different periods but also exhibiting credible intervals which do not cover negative values. To this end, we note that the optimal Black–Litterman portfolios are located further on the efficient frontier in comparison to the (objective) Bayesian optimal portfolio computed for the same value of the risk aversion coefficient and they also have larger prediction intervals indicating larger amount of uncertainty which is present when the optimal Black–Litterman portfolios are constructed.

## 5. Conclusion

The mean–variance analysis of Markowitz presents a fundamental method of portfolio construction which is very popular in the financial literature today. It provides an investor with the portfolio weights which determine the structure of the optimal portfolio. However, the investor faces a number of difficulties in implementing this procedure in practice. One of the main pitfalls of the mean–variance analysis is that its solution is presented in terms of unobservable quantities, the parameters of the asset returns distribution. As a result, the optimization problem is performed in two steps. After finding the analytical solution, the optimal portfolio is constructed by replacing the unknown parameters with their estimates. Due to the considerable influence of parameter uncertainty on the investment process, this procedure leads only to sub-optimal portfolios.

We deal with the problem from the viewpoint of Bayesian statistics. The optimization problem is formulated in terms of the posterior predictive distribution which does not involve unknown quantities. Consequently, we deal with parameter uncertainty before solving the optimization problem. This approach allows us to find optimal portfolio weights which now depend only on historical observations of the asset returns. The advantages of the approach are shown both theoretically and empirically. In particular, we show that the constructed Bayesian efficient frontier improves on the overoptimism present in the sample efficient frontier. Another important advantage of the suggested procedure is that it allows us not only to construct an optimal portfolio based on the posterior predictive distribution, but also leads to an intelligent technique by performing an interval forecast of future realizations of optimal portfolio returns which are obtained by employing the derived stochastic representation of the posterior predictive distribution.

## Acknowledgments

The authors would like to thank Professor Michael Dempster, Professor Jim Gatheral and two anonymous Reviewers for their helpful suggestions. This research was partly supported by the German Science Foundation (DFG) via the projects BO 3521/3-1 and SCHM 859/13-1 ‘Bayesian Estimation of

the Multi-Period Optimal Portfolio Weights and Risk Measures’ and by the Swedish Research Council (VR) via the project ‘Bayesian Analysis of Optimal Portfolios and Their Risk Measures’.

## Disclosure statement

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

## Funding

This research was partly supported by the German Science Foundation (DFG) via the projects BO 3521/3-1 and SCHM 859/13-1 ‘Bayesian Estimation of the Multi-Period Optimal Portfolio Weights and Risk Measures’ and by the Swedish Research Council (VR) via the project ‘Bayesian Analysis of Optimal Portfolios and Their Risk Measures’.

## ORCID

Nestor Parolya  <http://orcid.org/0000-0003-2147-2288>

## References

- Aguilar, O. and West, M., Bayesian dynamic factor models and portfolio allocation. *J. Bus. Econ. Stat.*, 2000, **18**, 338–357.
- Avramov, D. and Zhou, G., Bayesian portfolio analysis. *Annu. Rev. Financ. Econ.*, 2010, **2**, 25–47.
- Barry, C., Portfolio analysis under uncertain means, variances, and covariances. *J. Financ.*, 1974, **29**, 515–522.
- Basak, G.K., Jagannathan, R. and Ma, R., Estimating the risk in sample efficient portfolios. Technical report, Northwestern University, Evanston, IL, 2005.
- Bauder, D., Bodnar, R., Bodnar, T. and Schmid, W., Bayesian estimation of the efficient frontier. *Scand. J. Stat.*, 2019, **46**(3), 802–830.
- Bauder, D., Bodnar, T., Mazur, S., Okhrin, Y., *et al.*, Bayesian inference for the tangent portfolio. *Int. J. Theor. Appl. Financ.*, 2018, **21**(8), 1850054.
- Bauder, D., Bodnar, T., Parolya, N., Schmid, W., Bayesian inference of the multi-period optimal portfolio for an exponential utility. *J. Multivar. Anal.*, 2020, **175**, 104544.
- Bawa, V.S., Brown, S.J. and Klein, R.W., *Estimation Risk and Optimal Portfolio Choice*, 1979 (North-Holland Pub. Co: Amsterdam).
- Bernardo, J.M. and Smith, A.F.M., *Bayesian Theory*, 2000 (Wiley: Chichester).
- Best, M.J. and Grauer, R.R., On the sensitivity of mean–variance–efficient portfolios to changes in asset means: Some analytical and computational results. *Rev. Financ. Stud.*, 1991, **4**, 315–342.
- Black, F. and Litterman, R., Global portfolio optimization. *Financ. Analysts J.*, 1992, **48**, 28–43.
- Bodnar, O. and Bodnar, T., On the unbiased estimator of the efficient frontier. *Int. J. Theor. Appl. Financ.*, 2010, **13**, 1065–1073.
- Bodnar, T., Dmytriv, S., Parolya, N. and Schmid, W., Tests for the weights of the global minimum variance portfolio in a high-dimensional setting. *IEEE Trans. Signal Process.*, 2019, **67**(17), 4479–4493.
- Bodnar, T., Mazur, S. and Okhrin, Y., Bayesian estimation of the global minimum variance portfolio. *Eur. J. Oper. Res.*, 2017, **256**, 292–307.

Bodnar, T., Parolya, N. and Schmid, W., Estimation of the global minimum variance portfolio in high dimensions. *Eur. J. Oper. Res.*, 2018, **266**(1), 371–390.

Bodnar, T. and Schmid, W., Estimation of optimal portfolio compositions for gaussian returns. *Stat. Decis.*, 2008, **26**, 179–201.

Bodnar, T. and Schmid, W., Econometrical analysis of the sample efficient frontier. *Eur. J. Financ.*, 2009, **15**, 317–335.

Bradley, B.O. and Taqqu, M.S., Financial risk and heavy tails. In *Handbook of Heavy Tailed Distributions in Finance*, edited by S.T. Rachev, Volume 1 of *Handbooks in Finance*, pp. 35–103, 2003 (North-Holland: Amsterdam).

Brandt, M.W. and Santa-Clara, P., Dynamic portfolio selection by augmenting the asset space. *J. Financ.*, 2006, **61**(5), 2187–2217.

Broadie, M., Computing efficient frontiers using estimated parameters. *Ann. Oper. Res.*, 1993, **45**, 21–58.

Brown, S., Optimal portfolio choice under uncertainty: A Bayesian approach. PhD thesis, University of Chicago, 1976.

Frost, P. and Savarino, J., An empirical Bayes approach to efficient portfolio selection. *J. Financ. Quant. Anal.*, 1986, **21**, 293–305.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B., *Bayesian Data Analysis*, Vol. 2, 2014 (Chapman & Hall/CRC Boca Raton: FL, USA).

Givens, G.H. and Hoeting, J.A., *Computational Statistics*, 2012 (John Wiley & Sons: New York).

Gupta, A. and Nagar, D., *Matrix Variate Distributions*, 2000 (Chapman and Hall/CRC: Boca Raton).

Gupta, A., Varga, T. and Bodnar, T., *Elliptically Contoured Models in Statistics and Portfolio Theory*, 2nd ed., 2013 (Springer: New York).

Ingersoll, J.E., *Theory of Financial Decision Making*, 1987 (Rowman & Littlefield Publishers: Lanham, MD).

Jobson, J. and Korkie, B.M., Performance hypothesis testing with the Sharpe and Treynor measures. *J. Financ.*, 1981, **36**, 889–908.

Jorion, P., Bayes–Stein estimation for portfolio analysis. *J. Financ. Quant. Anal.*, 1986, **21**(3), 279–292.

Kacperczyk, M., Damien, P. and Walker, S.G., A new class of bayesian semi-parametric models with applications to option pricing. *Quant. Financ.*, 2013, **13**(6), 967–980.

Kacperczyk, M. and Damien, P., Asset allocation under distribution uncertainty. In *McCombs Research Paper Series No. IROM-01-11*, 2011 (Citeseer).

Kan, R. and Smith, D.R., The distribution of the sample minimum-variance frontier. *Manage. Sci.*, 2008, **54**, 1364–1380.

Kan, R. and Zhou, G., Optimal portfolio choice with parameter uncertainty. *J. Financ. Quant. Anal.*, 2007, **42**, 621–656.

Klein, R. and Bawa, V., The effect of estimation risk on optimal portfolio choice. *J. Financ. Econ.*, 1976, **3**, 215–231.

MacKinlay, A.C. and Pastor, L., Asset pricing models: Implications for expected returns and portfolio selection. *Rev. Financ. Stud.*, 2000, **13**, 883–916.

Markowitz, H., Portfolio selection. *J. Financ.*, 1952, **7**, 77–91.

Markowitz, H., *Portfolio Selection: Efficient Diversification of Investments*, 1959 (John Wiley: New York).

Merton, R.C., An analytic derivation of the efficient portfolio frontier. *J. Financ. Quant. Anal.*, 1972, **7**, 1851–1872.

Merton, R.C., On estimating the expected return on the market: An exploratory investigation. *J. Financ. Econ.*, 1980, **8**, 323–361.

Muirhead, R.J., *Aspects of Multivariate Statistical Theory*, 1982 (Wiley: New York).

Okhrin, Y. and Schmid, W., Distributional properties of portfolio weights. *J. Econom.*, 2006, **134**, 235–256.

Pástor, L., Portfolio selection and asset pricing models. *J. Financ.*, 2000, **55**, 179–223.

Pástor, L. and Stambaugh, R.F., Comparing asset pricing models: An investment perspective. *J. Financ. Econ.*, 2000, **56**, 335–381.

Rachev, S.T., Hsu, J.S.J., Bagasheva, B.S. and Fabozzi, F.J., *Bayesian Methods in Finance*, 2008 (Wiley: New Jersey).

Rousseeuw, P.J., Least median of squares regression. *J. Amer. Statist. Assoc.*, 1984, **79**(388), 871–880.

Rousseeuw, P.J. and Driessen, K.V., A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 1999, **41**(3), 212–223.

Sekerke, M., *Bayesian Risk Management: A Guide to Model Risk and Sequential Learning in Financial Markets*, 2015 (Wiley: New Jersey).

Siegel, A.F. and Woodgate, A., Performance of portfolios optimized with estimation error. *Manage. Sci.*, 2007, **53**, 1005–1015.

Stambaugh, R.F., Analyzing investments whose histories differ in length. *J. Financ. Econ.*, 1997, **45**, 285–331.

Tu, J. and Zhou, G., Incorporating economic objectives into bayesian priors: Portfolio choice under parameter uncertainty. *J. Financ. Quant. Anal.*, 2010, **45**, 959–986.

Würtl, D., Chalabi, Y., Chen, W. and Ellis, A., *Portfolio Optimization with R/Rmetrics*, 2015 (Rmetrics: Zurich).

## Appendix

*Proof of Theorem 1* The assumptions of infinitely exchangeability and multivariate centred spherically symmetry implies (see, e.g. Bernardo and Smith 2000, Proposition 4.6) that the asset returns are independently and identically distributed given the mean vector  $\mu$  and the covariance matrix  $\Sigma$  with the conditional distribution given by  $\mathbf{X}_t | \mu, \Sigma \sim \mathcal{N}_k(\mu, \Sigma)$  ( $k$ -dimensional normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$ ). Under this model with  $\theta = (\mu, \Sigma)$ , Jeffreys' prior is given by

$$\pi(\mu, \Sigma) \propto |\Sigma|^{-(k+1)/2}, \quad (\text{A1})$$

which leads to the posterior expressed as

$$\begin{aligned} \pi(\mu, \Sigma | \mathbf{x}_{(t-1)}) &\propto |\Sigma|^{-(n+k+1)/2} \exp \left\{ -\frac{n}{2} (\bar{\mathbf{x}}_{t-1} - \mu)^\top \right. \\ &\quad \left. \times \Sigma^{-1} (\bar{\mathbf{x}}_{t-1} - \mu) - \frac{1}{2} \text{tr}[\mathbf{S}_{t-1} \Sigma^{-1}] \right\}, \quad (\text{A2}) \end{aligned}$$

where  $\bar{\mathbf{x}}_{t-1}$  and  $\mathbf{S}_{t-1}$  are given in the statement of the theorem.

From (A2) we obtain that the posterior distribution of  $\Sigma$  is the inverse Wishart distribution (see Gupta and Nagar 2000 for the definition and properties) given by

$$\begin{aligned} \Sigma | \mu, \mathbf{x}_{t-1} &\sim \mathcal{IW}_k(n+k+1, \tilde{\mathbf{S}}_{t-1}(\mu)) \quad \text{with} \\ \tilde{\mathbf{S}}_{t-1}(\mu) &= \mathbf{S}_{t-1} + n(\mu - \bar{\mathbf{x}}_{t-1})(\mu - \bar{\mathbf{x}}_{t-1})^\top. \quad (\text{A3}) \end{aligned}$$

Furthermore, integrating out  $\Sigma$  we get the marginal posterior for  $\mu$  expressed as

$$\begin{aligned} \pi(\mu | \mathbf{x}_{(t-1)}) &\propto \int_{\Sigma > 0} |\Sigma|^{-(n+k+1)/2} \exp \left\{ -\frac{1}{2} \text{tr} \right. \\ &\quad \left. \times \left[ n(\bar{\mathbf{x}}_{t-1} - \mu)(\bar{\mathbf{x}}_{t-1} - \mu)^\top + \mathbf{S}_{t-1} \Sigma^{-1} \right] \right\} d\Sigma \\ &\propto |n(\bar{\mathbf{x}}_{t-1} - \mu)(\bar{\mathbf{x}}_{t-1} - \mu)^\top + \mathbf{S}_{t-1}|^{-n/2}, \end{aligned}$$

where the last equality follows by observing that the function under the integral is the density function of the inverse Wishart distribution with  $n+k+1$  degrees of freedom and parameter matrix  $n(\bar{\mathbf{x}}_{t-1} - \mu)(\bar{\mathbf{x}}_{t-1} - \mu)^\top + \mathbf{S}_{t-1}$ . The application of Sylvester's determinant theorem leads to

$$\pi(\mu | \mathbf{x}_{(t-1)}) \propto \left( 1 + n(\bar{\mathbf{x}}_{t-1} - \mu)^\top \mathbf{S}_{t-1}^{-1} (\bar{\mathbf{x}}_{t-1} - \mu) \right)^{-n/2}, \quad (\text{A4})$$

which proves that  $\mu | \mathbf{x}_{t-1} \sim t_k(n-k, \bar{\mathbf{x}}_{t-1}, (1/(n-k))\mathbf{S}_{t-1})$  ( $k$ -dimensional multivariate  $t$ -distribution with  $n-k$  degrees of freedom, location vector  $\bar{\mathbf{x}}_{t-1}$ , and scale matrix  $(1/(n-k))\mathbf{S}_{t-1}$ ).

Because  $\mathbf{X}_{t-n}, \dots, \mathbf{X}_t$  are independent given  $\mu$  and  $\Sigma$  as well as conditionally normally distributed, we get that the conditional distribution  $X_{p,t} | \mu, \Sigma$  coincides with  $X_{p,t} | \mu, \Sigma, \mathbf{x}_{(t-1)}$  is given by

$$X_{p,t} | \mu, \Sigma, \mathbf{x}_{(t-1)} \sim \mathcal{N}(\mathbf{w}^\top \mu, \mathbf{w}^\top \Sigma \mathbf{w}),$$

where the last equality proves that  $X_{p,t}$  depends on  $\mu$ ,  $\Sigma$ , and  $\mathbf{x}_{(t-1)}$  only over  $\mathbf{w}^\top \mu$  and  $\mathbf{w}^\top \Sigma \mathbf{w}$ .

The application of Theorem 3.2.13 in Muirhead (1982) leads to

$$\frac{\mathbf{w}^\top \Sigma \mathbf{w}}{\mathbf{w}^\top \tilde{\mathbf{S}}_{t-1}(\boldsymbol{\mu}) \mathbf{w}} \stackrel{d}{=} \frac{1}{\xi}, \quad (\text{A5})$$

where  $\xi \sim \chi_{n-k+1}^2$  and is independent of  $\boldsymbol{\mu}$  and  $\mathbf{X}_{(t-1)}$ . Then the stochastic representation of  $X_{p,t}$  is given by

$$X_{p,t} \stackrel{d}{=} \mathbf{w}^\top \boldsymbol{\mu} + \frac{\sqrt{\mathbf{w}^\top \tilde{\mathbf{S}}_{t-1}(\boldsymbol{\mu}) \mathbf{w}}}{\sqrt{n-k+1}} \tau_2,$$

where  $\tau_2 \sim t_1(n-k+1, 0, 1)$  is independent of  $\boldsymbol{\mu}$  and  $\mathbf{X}_{(t-1)}$ .

Finally, from the properties of the multivariate  $t$ -distribution, we obtain

$$\mathbf{w}^\top \boldsymbol{\mu} - \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} \sim t_1 \left( n-k, 0, \frac{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}}{n(n-k)} \right),$$

and, consequently,

$$X_{p,t} \stackrel{d}{=} \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}} \times \left( \frac{\tau_1}{\sqrt{n(n-k)}} + \sqrt{1 + \frac{\tau_1^2}{n-k}} \frac{\tau_2}{\sqrt{n-k+1}} \right),$$

where  $\tau_1$  and  $\tau_2$  are independent with  $\tau_1 \sim t_{n-k}$  and  $\tau_2 \sim t_{n-k+1}$ . ■

*Proof of Corollary 1.* In using the stochastic representation given in Theorem 1 and the properties of the  $t$ -distribution, we get

$$\begin{aligned} \mathbb{E}(\mathbf{X}_t | \mathbf{x}_{(t-1)}) &= \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}} \\ &\times \left( \frac{\mathbb{E}(\tau_1)}{\sqrt{n(n-k)}} + \mathbb{E} \left( \sqrt{1 + \frac{\tau_1^2}{n-k}} \right) \frac{\mathbb{E}(\tau_2)}{\sqrt{n-k+1}} \right) \\ &= \mathbf{w}^\top \bar{\mathbf{x}}_{t-1} \end{aligned}$$

and

$$\begin{aligned} \text{Var}(\mathbf{X}_t | \mathbf{x}_{(t-1)}) &= \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w} \text{Var} \left( \frac{\tau_1}{\sqrt{n(n-k)}} + \sqrt{1 + \frac{\tau_1^2}{n-k}} \frac{\tau_2}{\sqrt{n-k+1}} \right) \\ &= \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w} \left( \mathbb{E} \left( \frac{\tau_1^2}{n(n-k)} \right) + \mathbb{E} \left( \left( 1 + \frac{\tau_1^2}{n-k} \right) \frac{\tau_2^2}{n-k+1} \right) \right. \\ &\quad \left. + 2 \mathbb{E} \left( \frac{\tau_1}{\sqrt{n(n-k)}} \sqrt{1 + \frac{\tau_1^2}{n-k}} \frac{\tau_2}{\sqrt{n-k+1}} \right) \right) \end{aligned}$$

$$\begin{aligned} &= \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w} \left( \frac{1}{n(n-k)} \text{Var}(\tau_1) + \left( 1 + \frac{1}{n-k} \text{Var}(\tau_1) \right) \right. \\ &\quad \left. \times \frac{1}{n-k+1} \text{Var}(\tau_2) \right) \\ &= \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w} \left( \frac{1}{n(n-k)} \frac{n-k}{n-k-2} + \left( 1 + \frac{1}{n-k} \frac{n-k}{n-k-2} \right) \right. \\ &\quad \left. \times \frac{1}{n-k+1} \frac{n-k+1}{n-k-1} \right) \\ &= \left( \frac{1}{n-k-1} + \frac{2n-k-1}{n(n-k-1)(n-k-2)} \right) \mathbf{w}^\top \mathbf{S}_{t-1} \mathbf{w}. \end{aligned}$$

■

*Proof of Theorem 2.* Following the proof of Theorem 1, the marginal posterior of  $\boldsymbol{\mu}$  under the informative prior (16)–(17) is given by (see, e.g. Bauder *et al.* 2020, Proposition 2)

$$\boldsymbol{\mu} | \mathbf{x}_{t-1} \sim t_k \left( n + d_0 - 2k, \bar{\mathbf{x}}_{t-1,I}, \frac{1}{(n+r_0)(n+d_0-2k)} \mathbf{S}_{t-1,I} \right)$$

$$\text{with } \bar{\mathbf{x}}_{t-1,I} = \frac{n\bar{\mathbf{x}}_{t-1} + r_0 \mathbf{m}_0}{n+r_0} \quad \text{and}$$

$$\mathbf{S}_{t-1,I} = \mathbf{S}_{t-1} + \mathbf{S}_0 + nr_0 \frac{(\mathbf{m}_0 - \bar{\mathbf{x}}_{t-1,I})(\mathbf{m}_0 - \bar{\mathbf{x}}_{t-1,I})^\top}{n+r_0},$$

while the conditional posterior of  $\Sigma$  given  $\boldsymbol{\mu}$  is expressed as

$$\Sigma | \boldsymbol{\mu}, \mathbf{x}_{t-1} \sim IW_k(n + d_0 + 1, \mathbf{S}_{t-1,I}^*(\boldsymbol{\mu})) \quad \text{with}$$

$$\mathbf{S}_{t-1,I}^*(\boldsymbol{\mu}) = \mathbf{S}_{t-1,I} + (n+r_0)(\boldsymbol{\mu} - \bar{\mathbf{x}}_{t-1,I})(\boldsymbol{\mu} - \bar{\mathbf{x}}_{t-1,I})^\top.$$

Because  $\mathbf{X}_{t-n}, \dots, \mathbf{X}_t$  are independent given  $\boldsymbol{\mu}$  and  $\Sigma$  as well as conditionally normally distributed, we get that the conditional distribution  $X_{p,t} | \boldsymbol{\mu}, \Sigma$  coincides with  $X_{p,t} | \boldsymbol{\mu}, \Sigma, \mathbf{x}_{(t-1)}$  and is given by

$$X_{p,t} | \boldsymbol{\mu}, \Sigma, \mathbf{x}_{(t-1)} \sim \mathcal{N}(\mathbf{w}^\top \boldsymbol{\mu}, \mathbf{w}^\top \Sigma \mathbf{w}),$$

where the last equality proves that  $X_{p,t}$  depends on  $\boldsymbol{\mu}$ ,  $\Sigma$ , and  $\mathbf{x}_{(t-1)}$  only over  $\mathbf{w}^\top \boldsymbol{\mu}$  and  $\mathbf{w}^\top \Sigma \mathbf{w}$ .

From Muirhead (1982, Theorem 3.2.13) we get

$$\frac{\mathbf{w}^\top \Sigma \mathbf{w}}{\mathbf{w}^\top \mathbf{S}_{t-1,I}^*(\boldsymbol{\mu}) \mathbf{w}} \stackrel{d}{=} \frac{1}{\xi_I}, \quad (\text{A6})$$

where  $\xi_I \sim \chi_{n+d_0-2k+1}^2$  and is independent of  $\boldsymbol{\mu}$  and  $\mathbf{X}_{(t-1)}$ . Hence,

$$X_{p,t} \stackrel{d}{=} \mathbf{w}^\top \boldsymbol{\mu} + \frac{\sqrt{\mathbf{w}^\top \mathbf{S}_{t-1,I}^*(\boldsymbol{\mu}) \mathbf{w}}}{\sqrt{n+d_0-2k+1}} \eta_2,$$

where  $\eta_2 \sim t_1(n+d_0-2k+1, 0, 1)$  is independent of  $\boldsymbol{\mu}$  and  $\mathbf{X}_{(t-1)}$ .

Finally, from the properties of the multivariate  $t$ -distribution, we obtain

$$\mathbf{w}^\top \boldsymbol{\mu} - \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,I} \sim t_1 \left( n + d_0 - 2k, 0, \frac{\mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w}}{(n+r_0)(n+d_0-2k)} \right),$$

and, consequently,

$$\begin{aligned} X_{p,t} &\stackrel{d}{=} \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,I} + \sqrt{\mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w}} \left( \frac{\eta_1}{\sqrt{(n+r_0)(n+d_0-2k)}} \right. \\ &\quad \left. + \sqrt{1 + \frac{\eta_1^2}{n+d_0-2k}} \frac{\eta_2}{\sqrt{n+d_0-2k+1}} \right), \end{aligned}$$

where  $\eta_1$  and  $\eta_2$  are independent with  $\eta_1 \sim t_{n+d_0-2k}$  and  $\eta_2 \sim t_{n+d_0-2k+1}$ . ■

*Proof of Corollary 2.* The application of Theorem 2 leads to

$$\mathbb{E}(\mathbf{X}_t | \mathbf{x}_{(t-1)}) = \mathbf{w}^\top \bar{\mathbf{x}}_{t-1,I}$$

and

$$\begin{aligned} \mathbb{V}ar(\mathbf{X}_t | \mathbf{x}_{(t-1)}) &= \mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w} \\ &\quad \times \mathbb{V}ar \left( \frac{\eta_1}{\sqrt{(n+r_0)(n+d_0-2k)}} \right. \\ &\quad \left. + \sqrt{1 + \frac{\eta_1^2}{n+d_0-2k}} \frac{\eta_2}{\sqrt{n+d_0-2k+1}} \right) \\ &= \mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w} \left( \mathbb{E} \left( \frac{\eta_1^2}{(n+r_0)(n+d_0-2k)} \right) \right. \\ &\quad \left. + \mathbb{E} \left( \left( 1 + \frac{\tau_1^2}{n+d_0-2k} \right) \frac{\eta_2^2}{n+d_0-2k+1} \right) \right) \end{aligned}$$

$$\begin{aligned} &= \mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w} \left( \frac{\mathbb{V}ar(\eta_1)}{(n+r_0)(n+d_0-2k)} \right. \\ &\quad \left. + \left( 1 + \frac{\mathbb{V}ar(\eta_1)}{n+d_0-2k} \right) \frac{\mathbb{V}ar(\eta_2)}{n+d_0-2k+1} \right) \\ &= \mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w} \left( \frac{1}{(n+r_0)(n+d_0-2k-2)} \right. \\ &\quad \left. + \left( 1 + \frac{1}{n+d_0-2k-2} \right) \frac{1}{n+d_0-2k-1} \right) \\ &= \left( \frac{1}{n+d_0-2k-1} \right. \\ &\quad \left. + \frac{2n+r_0+d_0-2k-1}{(n+r_0)(n+d_0-2k-1)(n+d_0-2k-2)} \right) \\ &\quad \times \mathbf{w}^\top \mathbf{S}_{t-1,I} \mathbf{w}. \end{aligned}$$

■