Noname manuscript No.
(will be inserted by the editor)

Comparison and robustification of Bayes and Black-Litterman models

Katrin Schöttle · Ralf Werner · Rudi Zagst

Received: date / Accepted: date

Abstract For determining an optimal portfolio allocation, parameters representing the underlying market – characterized by expected asset returns and the covariance matrix – are needed. Traditionally, these point estimates for the parameters are obtained from historical data samples, but as experts often have strong opinions about (some of) these values, approaches to combine sample information and experts' views are sought for. The focus of this paper is on the two most popular of these frameworks – the Black-Litterman model and the Bayes approach. We will prove that – from the point of traditional portfolio optimization – the Black-Litterman is just a special case of the Bayes approach. In contrast to this, we will show that the extensions of both models to the robust portfolio framework yield two rather different robustified optimization problems.

Keywords portfolio optimization · robust optimization · Bayes model · Black-Litterman model

Katrin Schöttle
MEAG MUNICH ERGO AssetManagement GmbH, Oskar-von-Miller-Ring 18, 80333 München
Tel.: +49-89-24892062
Fax: +49-89-2489112062
E-mail: katrin.schoettle@gmx.de

Ralf Werner
Hypo Real Estate Holding AG, Group Risk Control, Unsöldstr. 2, 80538 München
Tel.: +49-89-203007263
Fax: +49-89-20300733263
E-mail: werner_ralf@gmx.net

Rudi Zagst
Zentrum Mathematik, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München
Tel.: +49-89-28917406
Fax: +49-89-28917407
E-mail: zagst@ma.tum.de

1 Introduction

The foundation of portfolio optimization was laid by Markowitz [19] in the 1950s when he introduced the mean-variance analysis which characterizes portfolios according to their expected return and volatility. Assuming a market with n risky assets and letting $\mu \in \mathbb{R}^n$ denote the vector of the expected asset returns and $\Sigma \in \mathbb{R}^{n \times n}$ their covariance matrix, the classical portfolio optimization[1] problem can be described as

$$\min_{x \in X} \ (1-\lambda)\sqrt{x^T \Sigma x} - \lambda x^T \mu \tag{P}$$

where the portfolio $x \in \mathbb{R}^n$ is given in terms of relative investments into the risky assets, i.e. it holds that $X \subset \{x \in \mathbb{R}^n \mid x^T \mathbf{1} = 1\}$. Naturally, the feasibility set X can contain additional constraints. The parameter $0 \leq \lambda \leq 1$ is used for tracing the whole efficient frontier.

As the market parameters μ and Σ are unknown, estimators $\hat{\mu}$ and $\hat{\Sigma}$ have to be determined – usually based on a historical data sample – and used instead for solving the optimization problem. It is widely known (see e.g. Jorion [17], Barry [1] or Best and Grauer [5]) that the solution x^* of problem (P) strongly depends on the particular values of $\hat{\mu}$ and $\hat{\Sigma}$ and that especially $\hat{\mu}$ has a significant influence on x^*. As furthermore the estimators largely depend on the underlying data sample, approaches to account for such parameter uncertainty are sought for, see e.g. the shrinkage approaches in Bawa, Brown and Klein [2]. An alternative possibility to obtain more robust optimal portfolios is the *robust counterpart approach* introduced by Ben-Tal and Nemirovski [3] in 1998. This approach is based on the idea to consider an entire set $\mathcal{U}$ of possible parameter realizations and to optimize the portfolio under the worst-case parameter. The robust counterpart problem to (P) is given by the following optimization problem:

$$\min_{x \in X} \max_{(r,C) \in \mathcal{U}} (1-\lambda)\sqrt{x^T C x} - \lambda x^T r. \tag{RP}$$

The arising question is how to define an appropriate uncertainty set $\mathcal{U}$ capturing the parameter uncertainty of the particular problem. Besides using a simple box constraint (e.g. $r \in [\hat{\mu} - \delta, \hat{\mu} + \delta]$) a straightforward definition is based on confidence ellipsoids. These two types of uncertainty sets are widely used both in the literature on robust portfolio optimization and for practical implementations, see e.g. Lutgens [18], Goldfarb and Iyengar [13], Ben-Tal, Margalit and Nemirovski [4], Cornuejols and Tütüncü [8], Meucci [20] or Schöttle and Werner [23]. Having the distribution of the uncertain parameter θ, which could be μ or Σ individually or the pair (μ, Σ), the confidence ellipsoid[2] is given as

$$\mathcal{U} = \left\{ \theta \mid (\theta - \mathbf{E}[\theta])^T \big(\mathbf{Cov}[\theta]\big)^{-1} (\theta - \mathbf{E}[\theta]) \leq \delta^2 \right\} \tag{1}$$

with the size δ^2 determined from the appropriate quantile expressing the desired level of confidence.

For actually solving the classical or the robust portfolio optimization problem, point estimates for $\mathbf{E}[\theta]$ and $\mathbf{Cov}[\theta]$, i.e. an explicit description of the uncertainty set,

---

[1] This formulation is completely equivalent to the classical textbook formulation, see e.g. Schöttle and Werner [24].

[2] See Meucci [20] or Schöttle and Werner [23] for details on how to transform the matrix Σ into a corresponding vector $\mathrm{vec}(\Sigma)$.

is necessary. Mostly, the *maximum likelihood estimators* are used to determine point estimates for μ and Σ based on a sample of (historical) observations. In case the sample of asset returns is supposed to follow an elliptical distribution, the distributions of the maximum likelihood estimators for μ and Σ are known as well (see Fang, Kotz and Ng [11]) which enables the creation of a confidence ellipsoid.

In many practical situations, experts have strong opinions about (some of) the market parameters which they want to use for determining their optimal portfolio. In addition to their view they do not want to neglect the historical performance of the assets under consideration, given in condensed form by the maximum likelihood estimator obtained from a data sample. Thus, approaches combining both market information and experts' opinions are gaining more and more interest. The general Bayesian method to calculate conditional distributions is rather well-known and discussed in many statistics books, see e.g. Press [22]. Using an empirical Bayesian approach to obtain more stable portfolio allocations is e.g. described in Jorion [16] or Frost and Savarino [12]. They both incorporate a prior related to the shrinkage approach of Stein, see e.g. Efron and Moris [10], i.e. assuming that all assets are identically distributed. The particular modification (shrinkage) of each asset return towards such a global average return is then determined by Bayesian methods.

Exploiting the resulting posterior distribution from the Bayesian approach to define an uncertainty set for the parameters in the robust portfolio optimization problem seems to be – in hindsight – a straightforward idea, but it was first presented in 2005 by Meucci [20]. Below, we recall the general Bayesian approach for determining point estimates for the parameters, and we also illustrate the robust Bayes approach of Meucci.

A rather different method to combine experts' opinions and market information was introduced by Black and Litterman in 1992, see [7]. There, a distributional assumption is made for the vector μ of expected asset returns which is then combined with both absolute and relative forecasts about the expected performance of individual assets. The covariance matrix which is also needed in portfolio opimization is assumed to be known and thus remains fix. Goldman Sachs – where the Black-Litterman approach was developed – has successfully applied the methodology and propagated the idea further in working papers in 1998 and 1999 ([6], [14]). Drobetz [9] analyzed the approach again from a practical point of view, and Idzorek [15] modified the model to incorporate user-specified confidence levels for the individual forecasts.

Our contribution to this area of research is threefold: We first illustrate how parameter uncertainty can be incorporated into the Black-Litterman model to obtain a robust version, similar to the existing robust Bayesian model. Further, we prove that the traditional Black-Litterman model is just a special case of the Bayesian model. Finally, we show that this is no longer true in the robust framework, i.e. the robust Bayesian model and the robust Black-Litterman model do not contain each other as special cases.

The rest of the paper is organized as follows: In Section 2 we present the (robust) Bayesian approach for determining point estimates as well as uncertainty sets. Section 3 discusses the Black-Litterman approach which originally yields a point estimate for μ, together with its extension for additionally obtaining an appropriate uncertainty set around μ. In Section 4 we compare the two frameworks mainly with respect to their generality, i.e. we investigate if one of the models contains the other one as a special case – both in the setting of obtaining point estimates for the classical portfolio optimization and when creating uncertainty sets for the robust problem.

2 The Bayesian approach

2.1 Prior assumption and market information

The general Bayesian approach is described e.g. in Press [22] and combines an externally given *prior assumption* about the distribution of the parameters and a sample of realizations, e.g. a sample of historical return data. As estimators for the parameters μ and Σ are sought for, a prior assumption about their joint distribution is necessary. We will assume that the pair of parameters (μ, Σ) follows a *normal inverse Wishart distribution*. This is a rather common assumption which is e.g. used in Meucci [20], [21], and usually made for mathematical convenience, as this is the natural conjugate prior distribution for the normal distribution, see Press [22]. The normal inverse Wishart distribution for a pair (μ, Σ) will be denoted by

$$(\mu, \Sigma) \sim \mathcal{NIW}(\mu_0, d_0, \Sigma_0, \nu_0) \tag{2}$$

and means that

$$\mu \mid \Sigma \sim \mathcal{N}\left(\mu_0, \frac{1}{d_0}\Sigma\right) \quad \text{and} \quad \Sigma \sim \mathcal{IW}(\nu_0 \Sigma_0, \nu_0 + n + 1)$$

where $\mathcal{IW}(A, p)$ denotes the *inverse Wishart distribution* with scale matrix A and p degrees of freedom, see e.g. Press [22].

Remark 1 For convenience of the reader, let us briefly review the main facts about the inverse Wishart distribution: It is closely related to the Wishart distribution by

$$\Sigma \sim \mathcal{IW}(\nu_0 \Sigma_0, \nu_0 + n + 1) \quad \Leftrightarrow \quad \Sigma^{-1} \sim \mathcal{W}\left(\frac{1}{\nu_0}\Sigma_0^{-1}, \nu_0\right).$$

The joint probability density function of (μ, Σ) is given by

$$\begin{aligned}
&\varphi_{\text{prior}}(\mu, \Sigma) = \varphi_{\mathcal{NIW}}(\mu, \Sigma) \\
&\quad = \gamma_1 |\Sigma^{-1}|^{\frac{\nu_0+n+2}{2}} \exp\left\{-\frac{1}{2}[d_0(\mu - \mu_0)^T \Sigma^{-1}(\mu - \mu_0) + \operatorname{tr}(\nu_0 \Sigma_0 \cdot \Sigma^{-1})]\right\}
\end{aligned}$$

with γ_1 denoting a normalizing constant such that $\varphi_{\text{prior}}(\mu, \Sigma)$ represents a probability density function.

Remark 2 In this setting, it obviously holds that

$$\mathbf{E}[\mu \mid \Sigma] = \mu_0, \quad \text{and} \quad \mathbf{Cov}[\mu \mid \Sigma] = \frac{1}{d_0}\Sigma.$$

In addition, as e.g. shown in Meucci [20], the marginal distribution of μ is given by the following Student-t distribution:

$$\mu \sim St\left(\mu_0, \frac{\nu_0}{\nu_0 - n + 1}\frac{1}{d_0}\Sigma_0, \nu_0 - n + 1\right)$$

and thus

$$\begin{aligned}
&\mathbf{E}[\mu] = \mu_0, \\
&\mathbf{Cov}[\mu] = \frac{\nu_0}{\nu_0 - n - 1}\frac{1}{d_0}\Sigma_0, \qquad \text{and, further} \\
&\mathbf{E}[\Sigma] = \frac{\nu_0}{\nu_0 - n - 1}\Sigma_0.
\end{aligned}$$

For given market parameters μ and Σ, it is assumed that the asset returns X are multivariate normally distributed, i.e. $X \mid \mu, \Sigma \sim \mathcal{N}(\mu, \Sigma)$. The market information is then given by i.i.d. realizations $x_1, \ldots, x_S$ of X. Their joint probability density function φ_M of $x_1, \ldots, x_S$ can hence be described by the product of the individual density function φ of x_s, $s = 1, \ldots, S$ and results in

$$\begin{aligned}\varphi_M(x_1, \ldots, x_S \mid \mu, \Sigma) &= \prod_{s=1}^{S} \varphi(x_s \mid \mu, \Sigma)\\ &= \left(\frac{1}{(2\pi)^{\frac{n}{2}}}\right)^S |\Sigma|^{-\frac{S}{2}} \exp\left\{-\frac{1}{2}\sum_{s=1}^{S}(x_s-\mu)^T \Sigma^{-1}(x_s-\mu)\right\}\\ &= \frac{1}{(2\pi)^{\frac{Sn}{2}}}|\Sigma|^{-\frac{S}{2}} \exp\left\{-\frac{1}{2}[S(\hat{\mu}-\mu)^T \Sigma^{-1}(\hat{\mu}-\mu) + \operatorname{tr}(S\hat{\Sigma}\Sigma^{-1})]\right\}.\end{aligned}$$

with

$$\hat{\mu} = \frac{1}{S}\sum_{s=1}^{S} x_s, \tag{3}$$

$$\hat{\Sigma} = \frac{1}{S}\sum_{s=1}^{S}(x_s-\hat{\mu})(x_s-\hat{\mu})^T. \tag{4}$$

It can be observed that the complete market information is summarized in only two variables, $\hat{\mu}$ and $\hat{\Sigma}$. Further note that $\hat{\mu}$ and $\hat{\Sigma}$ coincide with the maximum likelihood estimators for μ and Σ due to the multivariate normally distributed sample.

Remark 3 It can be shown (see e.g. Press [22]) that in the above setting, it holds that $\hat{\mu}$ is distributed independenly of $\hat{\Sigma}$ and that it actually holds

$$(\hat{\mu}, \hat{\Sigma}) \sim \mathcal{N}\big(\mu, \frac{1}{S}\Sigma\big) \times \mathcal{W}\left(\frac{1}{S}\Sigma, S-1\right). \tag{5}$$

2.2 Posterior distribution

Proposition 1 *In the Bayes setting defined by Equations (2) and (5), the* posterior distribution *of the parameters (μ, Σ) conditioned on the additional market information is given – in terms of its density*[3] *– by*

$$\begin{aligned}\varphi_{post}(\mu, \Sigma \mid x_1, \ldots, x_S) &= \gamma_2 \varphi_M(x_1, \ldots, x_S \mid \mu, \Sigma)\varphi_{prior}(\mu, \Sigma)\\ &= \gamma_2 \left(\frac{1}{(2\pi)^{\frac{n}{2}}}\right)^S |\Sigma|^{-\frac{S}{2}} \exp\left\{-\frac{1}{2}[S(\hat{\mu}-\mu)^T \Sigma^{-1}(\hat{\mu}-\mu) + tr(S\hat{\Sigma}\Sigma^{-1})]\right\}\\ &\quad \cdot \gamma_1 |\Sigma^{-1}|^{\frac{\nu_0+n+2}{2}} \exp\left\{-\frac{1}{2}[d_0(\mu-\mu_0)^T \Sigma^{-1}(\mu-\mu_0) + tr(\nu_0\Sigma_0 \cdot \Sigma^{-1})]\right\}\\ &= \gamma_3 |\Sigma|^{-\frac{\nu_1+n+2}{2}} \exp\left\{-\frac{1}{2}\left[d_1(\mu-\mu_1)^T \Sigma^{-1}(\mu-\mu_1) + tr(\nu_1\Sigma_1\Sigma^{-1})\right]\right\}\end{aligned}$$

[3] The parameters $\gamma_1, \gamma_2, \gamma_3$ denote the appropriate normalizing parameters to obtain probability density functions.

with the parameters

$$\begin{aligned}
\nu_1 &= \nu_0 + S, \\
d_1 &= d_0 + S, \\
\mu_1 &= \frac{d_0}{d_0 + S}\mu_0 + \frac{S}{d_0 + S}\hat{\mu}, \\
\Sigma_1 &= \frac{S}{\nu_0 + S}\hat{\Sigma} + \frac{\nu_0}{\nu_0 + S}\Sigma_0 + \frac{d_0}{d_0 + S}\frac{S}{\nu_0 + S}(\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T.
\end{aligned}$$

Corollary 1 *In the Bayes setting defined by Equations (2) and (5), the* posterior distribution *of the parameters* (μ, Σ) *conditioned on the additional market information is again a normal inverse Wishart distribution with adjusted parameters:*

$$\mu, \Sigma \mid x_1, \ldots, x_S \sim \mathcal{NIW}(\mu_1, d_1, \Sigma_1, \nu_1).$$

Proof This follows directly from looking at the density provided in Proposition 1.

Having an explicit posterior distribution allows both the determination of point estimates used in the classical portfolio optimization problem (Section 2.3) and the definition of an appropriate uncertainty set for the robust formulation (Section 2.4). For convenience, Table 1 summarizes the various distributions.

Table 1 Summary of the distributions in the Bayes model.

prior assumption	$(\mu, \Sigma) \sim \mathcal{NIW}(\mu_0, d_0, \Sigma_0, \nu_0)$
market distribution	$X \mid \mu, \Sigma \sim \mathcal{N}(\mu, \Sigma)$
posterior distribution	$\mu, \Sigma \mid x_1, \ldots, x_S \sim \mathcal{NIW}(\mu_1, d_1, \Sigma_1, \nu_1)$

Remark 4 Analogous to Remark 2, the marginal distribution of μ conditioned on the sample is given by a Student-t distribution with adjusted parameters:

$$\mu \mid x_1, \ldots, x_S \sim St\left(\mu_1, \frac{\nu_1}{\nu_1 - n + 1}\frac{1}{d_1}\Sigma_1, \nu_1 - n + 1\right)$$

2.3 Parameter point estimates

From now on we will work with the posterior distribution, as we are concerned with finding the appropriate optimization parameters for (P) and (RP). *Point estimates* for (P) for μ and Σ are obtained as expectation of the respective marginal distributions of $\mu \mid x_1, \ldots, x_S$ and $\Sigma \mid x_1, \ldots, x_S$. This is motivated by minimizing the expected loss, i.e. from solving

$$\hat{\theta} = \arg\min_{\zeta} \quad \mathbf{E}[\|\theta - \zeta\|_2^2], \tag{6}$$

which yields $\hat{\theta} = \mathbf{E}[\theta]$, with θ replacing μ or Σ in this formula. From Remarks 4 and 2, the according Bayesion point estimate $\hat{\mu}_B$ for μ can be calculated:

$$\hat{\mu}_B = \mathbf{E}[\mu \mid x_1, \ldots, x_S] \;=\; \mu_1 \;=\; \frac{d_0}{d_0 + S}\mu_0 + \frac{S}{d_0 + S}\hat{\mu}.$$

As the distribution of $\Sigma \mid x_1, \dots, x_S$ is given by an inverse Wishart distribution

$$\Sigma \mid x_1, \dots, x_S \sim \mathcal{IW}(\nu_1 \Sigma_1, \nu_1 + n + 1)$$

the expectation

$$\begin{aligned} \hat{\Sigma}_B = \mathbf{E}[\Sigma \mid x_1, \dots, x_S] &= \frac{\nu_1}{\nu_1 - n - 1} \Sigma_1 \\ &= \frac{1}{\nu_0 + S - n - 1} \left[S\hat{\Sigma} + \nu_0 \Sigma_0 + \frac{d_0 S}{d_0 + S} (\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T \right]. \end{aligned}$$

can be derived according to Remark 1.

2.4 Incorporating uncertainty

Based on the first and second moments of the posterior distributions of the unknown parameters, an uncertainty set for (RP) can be defined as in Equation (1). For two reasons, we here focus on the setting of considering uncertainty only in the parameter μ and assume that the covariance matrix is known or appropriately estimated by $\hat{\Sigma}_B$, e.g. from historical data. First, the parameter μ is influencing the result of the optimization problem rather strongly; small changes in μ alter the optimal solution significantly more than slight modifications in Σ. Second, the handling of uncertainty sets for matrices is rather involved and hence avoided in practical settings.

Considering an uncertainty set for μ in the Bayes model, the confidence ellipsoid based on the first two moments of the marginal posterior distribution of $\mu \mid x_1, \dots, x_S$ is then – according to Equation (1) – defined as follows:

$$\mathcal{U}_B = \left\{ r \in \mathbb{R}^n \mid (r - m_B)^T C_B^{-1} (r - m_B) \leq \delta^2 \right\}$$

with

$$\begin{aligned} m_B &= \mathbf{E}[\mu \mid x_1, \dots, x_S] = \mu_1 = \frac{d_0}{d_0 + S} \mu_0 + \frac{S}{d_0 + S} \hat{\mu} = \hat{\mu}_B, \\ C_B &= \mathbf{Cov}[\mu \mid x_1, \dots, x_S] = \frac{\nu_1}{\nu_1 - n - 1} \frac{1}{d_1} \Sigma_1 = \frac{1}{d_1} \hat{\Sigma}_B. \end{aligned}$$

Note that the midpoint m_B of the ellipsoidal uncertainty set always coincides with the point estimate $\hat{\mu}_B$ due to construction. Further, in this case it holds that $C_B = \frac{1}{d_1} \hat{\Sigma}_B$, as expected from the ellipsoidal distribution structure.

Remark 5 As already mentioned above, in the Bayesian setting both μ and Σ are exposed to uncertainty. Thus, there are further possibilities to perform a robust portfolio optimization when considering uncertainty around Σ as well, as analyzed e.g. in Meucci [20] or in Schöttle and Werner [24]. In this case, $\mathbf{Cov}[\Sigma \mid x_1, \dots, x_S]$ needs to be calculated. Although this is technically involved, it is doable as the moments of the (inverse) Wishart distribution are known.

3 Black-Litterman approach

3.1 Prior assumption and forecasts

A second approach combining expert opinion and market information to a new parameter estimate for the return vector is the *Black-Litterman model*, developed by Black and Litterman [7]. In the Black-Litterman setting, only the return vector μ is explicitly exposed to uncertainty, the covariance matrix Σ is supposed to be known, i.e. estimated by $\hat{\Sigma}$ without uncertainty. As the result of this approach is generally the entire *distribution* of μ and not only a point estimate, the Black-Litterman model can naturally be used for the definition of an uncertainty set as well. In our setting, the *prior assumption* is given by

$$\mu \sim \mathcal{N}(\hat{\mu}, \tau \frac{1}{S} \hat{\Sigma})$$

with $\hat{\mu}$ and $\hat{\Sigma}$ as in formulas (3) and (4). Note that in contrast to the Bayesian approach, the market information, i.e. the sample data $x_1, \ldots, x_S$, completely determines the prior.

Remark 6 In the original work by Black and Litterman, the return estimate $\hat{\mu}$ is deduced from the global market portfolio by equilibrium considerations, whereas in our setting, it is based on the maximum likelihood estimator. However, the source of the return estimate does not matter and the subsequent analysis would remain the same for any other choice for $\hat{\mu}$. In accordance with the original framework, the covariance of the return estimate is set proportional to $\hat{\Sigma}$.

Instead of relying on the traditional framework presented in [7], we prefer to generalize the setting to a Student-t distribution[4] with k ($k \geq 3$) degrees of freedom,

$$\mu \sim St(\hat{\mu}, \tau \hat{\Sigma}_k, k), \quad \text{with } \hat{\Sigma}_k = \frac{1}{S} \frac{k-2}{k} \hat{\Sigma}, \tag{7}$$

which we call the *extended Black-Litterman framework*. Obviously, if we set the degrees of freedom $k = \infty$, we can recover the *original Black-Litterman approach*. Note that the prior distribution of μ belongs to the same class as the posterior distribution in the Bayes setting, see Section 2.2.

Remark 7 As $\mu \sim St(\hat{\mu}, \tau \hat{\Sigma}_k, k)$ it holds that $\mathbf{E}[\mu] = \hat{\mu}$ and $\mathbf{Cov}[\mu] = \tau \frac{1}{S} \hat{\Sigma}$, independent of k.

As the covariance matrix Σ is assumed to be known, investor forecasts can only be made for the uncertain vector μ. Such forecasts can be any absolute or relative opinion about individual assets which are mathematically expressed in the form

$$Q = P\mu + \varepsilon \tag{8}$$

[4] Note that the scaling factor $\frac{k-2}{k}$ has to be introduced due to the fatter tails of the Student-t distribution. For more details on the multivariate Student-t distribution, we refer to Press [22] or Meucci [20].

with the *forecast* $Q \in \mathbb{R}^m$ and the *pick matrix* $P \in \mathbb{R}^{m \times n}$. Here, ε and μ are assumed to be jointly Student-t distributed[5] with the same degrees of freedom k, $\mathbf{Cov}[\varepsilon] = (1-\tau)\Omega_k$ and $\mathbf{Cov}[\varepsilon, \mu] = 0$:

$$\begin{pmatrix} \mu \\ \varepsilon \end{pmatrix} \sim St\left(\begin{pmatrix} \hat{\mu} \\ 0 \end{pmatrix}, \begin{pmatrix} \tau \hat{\Sigma}_k & 0 \\ 0 & (1-\tau)\Omega_k \end{pmatrix}, k \right). \tag{9}$$

The forecast vector Q consists of the explicit forecast values, and the pick matrix P contains the information which assets are affected by the respective forecasts. Without loss of generality, one can assume that P has full rank m ($m \leq n$), i.e. one can presume that P does not contain any redundant forecast.

Example 1 Let us consider three assets A, B and C, together with the *relative forecast* "A outperforms B on average by 4%" and the *absolute forecast* "C has an expected return of 6%". In this case, Q and P would be given by

$$P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad Q = \begin{pmatrix} 4\% \\ 6\% \end{pmatrix}.$$

Remark 8 For given μ, for the forecast Q it obviously holds that

$$Q \mid \mu \sim St(P\mu, (1-\tau)\Omega_k, k),$$

i.e. the matrix $(1-\tau)\Omega_k$ describes the variance of Q (given μ) and thus expresses the confidence about the individual forecasts. The parameter $\tau \in [0,1]$ can be used to weight between the uncertainty of the market and the uncertainty of the forecast. High confidence in the prior, i.e. in the market information, corresponds to $\tau = 0$, whereas $\tau = 1$ represents high confidence in the forecast, resp. In this aspect, our choice of τ differs from the existing literature, as we neither suggest small nor large τ. Please note that in some of these approaches (e.g. He and Litterman, [14]), the factor $\frac{1}{S}$ is already included in τ, which would result in rather small values for τ.

Remark 9 As P has full rank, each $\varepsilon \in \mathbb{R}^m$ can be expressed as $\varepsilon = P\xi$, with $\xi \in \mathbb{R}^n$. For example, ξ could be set as

$$\xi = \arg \min_{\substack{\rho \in \mathbb{R}^n \\ P\rho = \varepsilon}} ||\rho||_2^2, \quad \text{i.e.} \quad \xi = P^T(PP^T)^{-1}\varepsilon. \tag{10}$$

This means that we can suppose that the uncertainty ε in the forecast Q is due to an *lifted uncertainty* ξ in the expected return μ, i.e.

$$Q = P(\mu + \xi).$$

This corresponds to the situation that the forecast is made on a disturbed return estimate μ with disturbance ξ. Therefore, it is quite natural to assume that ξ has the same uncertainty structure as μ: instead of the general form (9), one then assumes more specifically that

$$\begin{pmatrix} \mu \\ \xi \end{pmatrix} \sim St\left(\begin{pmatrix} \hat{\mu} \\ 0 \end{pmatrix}, \begin{pmatrix} \tau \hat{\Sigma}_k & 0 \\ 0 & (1-\tau)\hat{\Sigma}_k \end{pmatrix}, k \right). \tag{11}$$

or, equivalently, $\Omega_k = P\hat{\Sigma}_k P^T$. A very similar assumption can e.g. be found in Idzorek [15], at least for the diagonal elements of Ω, although other authors, like He and Litterman [14] prefer alternative specifications.

5 In this setting, ε and μ are uncorrelated, but not independent. Only in the limiting case of a normal distribution ($k = \infty$), we obtain independence.

3.2 Posterior distribution

From Equations (8) and (9) we can deduce the joint distribution of μ and Q as well as the conditional distribution of μ given $Q = q$.

Proposition 2 *In the extended Black-Litterman framework defined by Equations (8) and (9) it holds:*

(i) The joint distribution of μ and Q is given by

$$\begin{pmatrix} \mu \\ Q \end{pmatrix} \sim St\left(\begin{pmatrix} \hat{\mu} \\ P\hat{\mu} \end{pmatrix}, \begin{pmatrix} \tau\hat{\Sigma}_k & \tau\hat{\Sigma}_k P^T \\ \tau P\hat{\Sigma}_k & V_k \end{pmatrix}, k \right),$$

where $V_k = (1-\tau)\Omega_k + \tau P\hat{\Sigma}_k P^T$.

(ii) The conditional distribution of $\mu \mid Q = q$ is given by

$$\mu \mid Q = q \sim St\big(\hat{\mu} + \tau\hat{\Sigma}_k P^T V_k^{-1}(q - P\hat{\mu}), \tau\hat{\Sigma}_k - \tau\hat{\Sigma}_k P^T V_k^{-1}\tau P\hat{\Sigma}_k, k+n\big).$$

Proof

(i) Expressing

$$\begin{pmatrix} \mu \\ Q \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ P & I_m \end{pmatrix} \begin{pmatrix} \mu \\ \varepsilon \end{pmatrix}$$

it straightforwardly holds that the pair (μ, Q) is again Student-t distributed, see e.g. Fang, Kotz and Ng [11], Theorem 2.16:

$$\begin{aligned} \begin{pmatrix} \mu \\ Q \end{pmatrix} &\sim St\left(\begin{pmatrix} I_n & 0 \\ P & I_m \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ 0 \end{pmatrix}, \begin{pmatrix} I_n & 0 \\ P & I_m \end{pmatrix} \begin{pmatrix} \tau\hat{\Sigma}_k & 0 \\ 0 & (1-\tau)\Omega_k \end{pmatrix} \begin{pmatrix} I_n & 0 \\ P & I_m \end{pmatrix}^T, k \right) \\ &= St\left(\begin{pmatrix} \hat{\mu} \\ P\hat{\mu} \end{pmatrix}, \begin{pmatrix} \tau\hat{\Sigma}_k & \tau\hat{\Sigma}_k P^T \\ \tau P\hat{\Sigma}_k & (1-\tau)\Omega_k + \tau P\hat{\Sigma}_k P^T \end{pmatrix}, k \right). \end{aligned}$$

(ii) Consequence of the conditioning formula for elliptical distribution, see e.g. Fang Kotz and Ng [11], Theorem 2.18.

For better readability, Table 2 summarizes the different distributions within the extended Black-Litterman setting.

Table 2 Summary of the distributions in the extended Black-Litterman model with general V_k.

prior assumption	$\mu \sim St(\hat{\mu}, \tau\hat{\Sigma}_k, k)$, $\Sigma = \hat{\Sigma}$
forecasts	$Q = P\mu + \varepsilon \quad \Rightarrow \quad Q \sim St(P\hat{\mu}, V_k, k)$
posterior distribution	$\mu \mid Q \sim St\big(\hat{\mu} + \tau\hat{\Sigma}_k P^T V_k^{-1}(q - P\hat{\mu}), \tau\hat{\Sigma}_k - \tau\hat{\Sigma}_k P^T V_k^{-1}\tau P\hat{\Sigma}_k, k+n\big)$

Corollary 2 *Replacing assumption (9) by assumption (11), we have $V_k = P\hat{\Sigma}_k P^T$ instead of $V_k = (1-\tau)\Omega_k + \tau P\hat{\Sigma}_k P^T$ in Proposition 2.*

From now on, we will assume that (11) holds instead of (9), i.e. $V_k = P\hat{\Sigma}_k P^T$.

3.3 Parameter point estimates

According to Equation (6), the *point estimates* for μ and Σ in the extended Black-Litterman model are given by the expectation of the posterior distribution $\mu \mid Q$ and $\hat{\Sigma}$, respectively, as no uncertainty was assumed for the covariance matrix:

$$\hat{\mu}_{BL} = \mathbf{E}[\mu|Q = q] = \hat{\mu} + \tau\hat{\Sigma}P^T(P\hat{\Sigma}P^T)^{-1}(q - P\hat{\mu}) \tag{12}$$

$$\hat{\Sigma}_{BL} = \hat{\Sigma}. \tag{13}$$

Note that all factors of $\hat{\Sigma}$ in the definition of $\hat{\Sigma}_k$ cancel out in (12) and thus $\hat{\mu}_{BL}$ is independent of k. In other words, we recover exactly the same result as in the original Black-Litterman framework, independent of k. The parameter k only enters the calculations later on, when we consider uncertainty sets.

Remark 10 One of the more frequent usages of the Black-Litterman model is the case of absolute forecasts for each asset, i.e. $P = I$. In this case, Equation (12) simplifies to

$$\hat{\mu}_{BL} = \tau q + (1 - \tau)\hat{\mu}.$$

Hence, the resulting estimate $\hat{\mu}_{BL}$ is a convex combination of the prior $\hat{\mu}$ and the forecast q and the trade-off between these two is controlled by the parameter τ.

Remark 11 As P has full rank, the forecast q can e.g. be written as $q = Pq_l$, where q_l is the *lifted forecast* defined as $q_l = P^T(PP^T)^{-1}q$, similar to (10). Based on this lifted forecast, the above Equation (12) can be rewritten as:

$$\begin{aligned}
\hat{\mu}_{BL} &= \hat{\mu} + \tau\hat{\Sigma}P^T(P\hat{\Sigma}P^T)^{-1}(Pq_l - P\hat{\mu}) \\
&= \hat{\mu} + \tau\underbrace{\hat{\Sigma}P^T(P\hat{\Sigma}P^T)^{-1}P}_{=:K}(q_l - \hat{\mu}) \qquad (*) \\
&= \hat{\mu} + \tau K(q_l - \hat{\mu}) \\
&= (I - \tau K)\hat{\mu} + \tau K q_l.
\end{aligned}$$

It is easy to see that K is a projection matrix, i.e. $K^2 = K$ and $K(I - K) = 0$. Multiplying the above equation for $\hat{\mu}_{BL}$ with K and $I - K$, resp., we obtain

$$\begin{aligned}
K\hat{\mu}_{BL} &= (1 - \tau)K\hat{\mu} + \tau K q_l, \quad \text{and} \\
(I - K)\hat{\mu}_{BL} &= (I - K)\hat{\mu}.
\end{aligned}$$

These two equations show that in the range of K, the Black-Litterman estimator $\hat{\mu}_{BL}$ is a convex combination of $\hat{\mu}$ and the lifted forecast q_l, whereas in the kernel of K, i.e. in the range of $I - K$, the Black-Litterman estimator coincides with $\hat{\mu}$. Combining these two facts, this can be nicely interpreted: the Black-Litterman estimator keeps $\hat{\mu}$ where no forecast is made, and averages out between $\hat{\mu}$ and the (lifted) forecast q_l, where a forecast is made.

3.4 Incorporating uncertainty

Apart from obtaining Black-Litterman point estimates, the posterior distribution also allows the creation of an uncertainty set for the parameter μ. Analogous to the Bayesian setting, the first two moments are used to define the confidence ellipsoid

$$\mathcal{U}_{BL} = \left\{ r \in \mathbb{R}^n \mid (r - m_{BL})^T C_{BL}^{-1} (r - m_{BL}) \leq \delta^2 \right\}$$

with

$$\begin{aligned} m_{BL} &= \mathbf{E}[\mu | Q = q] = \hat{\mu} + \tau K(q_l - \hat{\mu}) = \hat{\mu}_{BL}, \\ C_{BL} &= \mathbf{Cov}[\mu | Q = q] = \tau(I - \tau K)\hat{\Sigma}_k. \end{aligned}$$

For $\tau = 0$ we get $C_{BL} = 0$, whereas for $\tau = 1$ one obtains $C_{BL} = (I - K)\hat{\Sigma}_k$. This means that in case of a 100% confidence in the prior $\hat{\mu}$, no uncertainty around this estimate is to be expected. In contrast to this, in case of absolute certainty on the forecast (i.e. $\tau = 1$), the uncertainty $\hat{\Sigma}$ is projected onto (the space of) those assets, where no forecast is made. If forecasts for all assets are made, i.e. P has full rank, then $K = I$ and $C_{BL} = 0$. It is interesting to note that the uncertainty in the Black-Litterman setting is always smaller[6] than in the traditional robust optimization setting – which is due to the fact that the additional information of the forecast always reduces the uncertainty. Quite obviously, the highest uncertainty is obtained for $\tau = \frac{1}{2}$, when neither the prior nor the forecast is trustworthy.

4 Comparison of the Bayes and the extended Black-Litterman model

Both the Bayesian and the Black-Litterman approach are models combining market and expert information, but they do it in different ways:

- In the Bayes model the expert first describes the prior distribution of the parameters which is then conditioned by the market data.
- In the Black-Litterman model the prior is determined by the data sample, and the experts' opinions alter the outcome.

Furthermore, it holds that both models can be applied – as illustrated above – in two different ways:

- Use the point estimates for μ and Σ to solve the classical portfolio optimization problem.
- Determine an uncertainty set for μ from the according posterior distribution and solve the robust portfolio optimization problem.

Judging from the setup of the models, none seems to contain the other one as a special case. The Bayesian setting has some flexibility in modeling the covariance matrix, whereas in the Black-Litterman model the covariance matrix is fixed. On the other hand, the Black-Litterman model allows more possibilities for incorporating experts' opinions. Both absolute and relative forecasts can be made on arbitrary assets. In the

[6] Using the Löwner ordering of matrices, it can be shown that $\tau(I - \tau K)\hat{\Sigma}_k \preceq \frac{1}{S}\hat{\Sigma}$, which can be interpreted as smaller uncertainty. Here, the notation $A \preceq B$ means that $B - A$ is a positive semidefinite matrix.

Bayes model however, external knowledge only enters the model through the prior assumption which leads to the impression that only an absolute return forecast for each individual asset can be made. We will show in Theorem 1 that this impression is misleading, at least if the traditional portfolio optimization framework is considered.

Subsequently, we first compare the respective *point estimates* which are needed for the classical optimization problem, afterwards, we investigate the match of the *uncertainty sets* for the robust portfolio optimization. Before starting the comparison of parameter point estimates or uncertainty sets, we summarize all the considered parameters in Table 3.

Table 3 Summary of the parameters in the Bayes and extended Black-Litterman model.

	Bayes	extended Black-Litterman
μ	$\hat{\mu}_B = \frac{d_0}{d_0+S}\mu_0 + \frac{S}{d_0+S}\hat{\mu}$	$\hat{\mu}_{BL} = \hat{\mu} + \tau K(q_l - \hat{\mu})$
Σ	$\hat{\Sigma}_B = \frac{1}{\nu_0+S-n-1}\left[S\hat{\Sigma} + \nu_0\Sigma_0 + \frac{d_0 S}{d_0+S}(\mu_0-\hat{\mu})(\mu_0-\hat{\mu})^T\right]$	$\hat{\Sigma}_{BL} = \hat{\Sigma}$
m	$m_B = \frac{d_0}{d_0+S}\mu_0 + \frac{S}{d_0+S}\hat{\mu}$	$m_{BL} = \hat{\mu} + \tau K(q_l - \hat{\mu})$
C	$C_B = \frac{1}{\nu_0+S-n-1} \cdot \frac{1}{d_0+S}\left[S\hat{\Sigma} + \nu_0\Sigma_0 + \frac{d_0 S}{d_0+S}(\mu_0-\hat{\mu})(\mu_0-\hat{\mu})^T\right]$	$C_{BL} = \tau(I - \tau K)\hat{\Sigma}_k$

4.1 Comparing point estimates

We first compare the point estimates of the Bayes model and the extended Black-Litterman approach. Obviously, in the Black-Litterman model, $\hat{\Sigma}_{BL}$ is fixed and completely determined by the data. Hence, the Black-Litterman model cannot be as flexible as the Bayes model. The opposite direction will be shown in the following theorem, i.e. the Bayes model can be parametrized in such a way that arbitrary Black-Litterman parameters can be reproduced.

Theorem 1 *Let the Black-Litterman model, i.e. parameters P, q, τ and k, be given. Then, there exists a $M \in \mathbb{R}_+$ such that the following choice of parameters for the Bayes model reproduces the same point estimates as the Black-Litterman model, i.e. $\hat{\mu}_B = \hat{\mu}_{BL}$ and $\hat{\Sigma}_B = \hat{\Sigma}_{BL}$:*

$$\begin{aligned} d_0 &= \frac{\tau}{1-\tau}S \\ \nu_0 &> M \text{ arbitrarily} \\ \mu_0 &= Kq_l + (I-K)\hat{\mu} \\ \Sigma_0 &= \hat{\Sigma} - \frac{n+1}{\nu_0}\hat{\Sigma} - \frac{S\tau}{\nu_0}(\mu_0-\hat{\mu})(\mu_0-\hat{\mu})^T. \end{aligned}$$

Proof Let us consider the equations $\hat{\mu}_B = \hat{\mu}_{BL}$ and $\hat{\Sigma}_B = \hat{\Sigma}_{BL}$ in more detail:

$$\text{(E1)} \quad \frac{d_0}{d_0+S}\mu_0 + \frac{S}{d_0+S}\hat{\mu} = \hat{\mu} + \tau K(q_l - \hat{\mu}),$$

$$\text{(E2)} \quad S\hat{\Sigma} + \nu_0\Sigma_0 + \frac{d_0 S}{d_0+S}(\mu_0-\hat{\mu})(\mu_0-\hat{\mu})^T \;=\; (\nu_0+S-n-1)\hat{\Sigma}.$$

If we set $d_0 = \frac{\tau}{1-\tau}S$, then (E1) can be solved for μ_0, which yields $\mu_0 = Kq_l + (I-K)\hat{\mu}$. Plugging d_0 in (E2) and substracting $S\hat{\Sigma}$ from both sides, (E2) becomes

$$\nu_0\Sigma_0 + S\tau(\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T \;=\; (\nu_0 - n - 1)\hat{\Sigma}.$$

If we divide by ν_0 and solve for Σ_0 we get:

$$\Sigma_0 \;=\; \hat{\Sigma} - \frac{n+1}{\nu_0}\hat{\Sigma} - \frac{S\tau}{\nu_0}(\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T.$$

As $\hat{\Sigma}$ is positive definite, it is possible to choose ν_0 large enough such that the right hand side still remains a positive definite matrix. We set M to the lowest ν_0 for which the right hand side remains positive semidefinite, i.e.

$$\begin{aligned} M = \min_{\nu \in R} \;& \nu \\ & \text{subject to} \\ & \nu\hat{\Sigma} \succeq (n+1)\hat{\Sigma} + S\tau(\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T. \end{aligned}$$

which proves the claim. □

Remark 12 Eventually, d_0 was chosen as $\frac{\tau}{1-\tau}S$ for convenience, as this leads to a rather intuitive expression for μ_0. However, any other choice of d_0 is feasible as well. Depending on d_0, different choices of μ_0 and M would then be necessary. We will see that this additional flexibility is actually needed in the next section.

Remark 13 From the above Theorem 1 it can be seen that by expressing Black-Litterman forecasts within the sound statistical method of Bayes, the Bayesian approach is not restricted to giving only absolute forecasts. Expressing relative forecasts in the prior is as well possible by making use of the Black-Litterman idea, even though this possibility has not been obvious from the setup of the Bayes model. Following similar arguments as in the previous section, it can be deduced that μ_0 coincides with the lifted forecast q_l in the range of K, whereas it is equal to $\hat{\mu}$ in its null space. If $K = I$, then it actually holds that $\mu_0 = q_l$.

Remark 14 If we let $\nu_0 = \infty$ in Theorem 1, the choice for Σ_0 simplifies to $\Sigma_0 = \hat{\Sigma}$. In this case, the covariance matrix in the Bayes setting is not subject to any uncertainty, but fixed to $\hat{\Sigma}$ as well.

4.2 Comparing uncertainty sets

After comparing the point estimates for μ and Σ, we next compare the parameters for the robust portfolio optimization. As before, we only account for uncertainty around μ explicitly. Hence, for a comparison of the robust settings, we need to match the uncertainty sets for the return vector (given in terms of the midpoint and the shape matrix) and the point estimate for the covariance matrix, i.e. it has to hold

$$m_B = m_{BL}, \quad C_B = C_{BL} \quad \text{and} \quad \hat{\Sigma}_B = \hat{\Sigma}_{BL}.$$

As mentioned, $m_B = m_{BL}$ is equivalent to $\mu_B = \mu_{BL}$. This means that the first and the last equality were already investigated in the setting of the previous theorem.

Therefore, the same observation as before holds true: the robust Black-Litterman setting is not as flexible as the robust Bayes setting and corresponding Bayes estimates cannot always be reproduced in the Black-Litterman setting. In addition, if we have a closer look at the connection between the covariance of μ and the expectation of Σ, we see that

$$C_B = \frac{1}{d_0 + S}\hat{\Sigma}_B \quad \text{and} \quad C_{BL} = \tau(I - \tau K)\frac{1}{S}\frac{k-2}{k}\hat{\Sigma}_{BL}.$$

If $\hat{\Sigma}_B = \hat{\Sigma}_{BL}$ holds, see Theorem 1, and keeping in mind that K needs to be a projection matrix, then the above equation can be fulfilled if and only if $K = I$. Otherwise, neither the Bayes nor the Black-Litterman setting contains the other as a special case. This is a very remarkable conclusion, as in the traditional framework the Bayes setting dominates the Black-Litterman setting.

Remark 15 As shown in Schöttle and Werner [24], the shape of the robust efficient frontier depends on the fact that C_B is a multiple of $\hat{\Sigma}_B$, which is not the case in the Black-Litterman framework. Thus, in the Bayes framework, the robust frontier coincides with the left part of the traditional Bayesian efficient frontier in contrast to the Black-Litterman setting where the robust frontier may take different shapes. This difference can be used to judge which of the robustified frameworks is favored in practical applications.

Remark 16 If a full set of n forecasts is made in the Black-Litterman setting, then P has full rank n. This means that P is invertible and thus $K = I$. In this case, Theorem 1 can be generalized to the robust setting, using the additional flexibility in the choice of d_0.

Theorem 2 *Let the Black-Litterman model, i.e. parameters P, q, τ and k, be given and let $K = I$ (i.e. rank$(P) = n$). Then there exists an $M \in \mathbb{R}_+$ such that the following choice of parameters for the robust Bayes model reproduces the same point estimates – and thus the same uncertainty sets – as the robust Black-Litterman model, i.e. $\hat{\mu}_B = \hat{\mu}_{BL}$, $\hat{\Sigma}_B = \hat{\Sigma}_{BL}$ and $C_B = C_{BL}$:*

$$\begin{aligned}
d_0 &= \Big(\frac{1}{\tau(1-\tau)}\frac{k}{k-2} - 1\Big)S \\
\nu_0 &> M \ \textit{arbitrarily} \\
\mu_0 &= \hat{\mu} + \frac{d_0 + S}{d_0}\tau(q_l - \hat{\mu}) \\
\Sigma_0 &= \hat{\Sigma} - \frac{n+1}{\nu_0}\hat{\Sigma} - \frac{S\tau}{\nu_0}(\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T.
\end{aligned}$$

Proof Let us again consider the equations $\hat{\mu}_B = \hat{\mu}_{BL}$ and $\hat{\Sigma}_B = \hat{\Sigma}_{BL}$ as well as $C_B = C_{BL}$:

$$\begin{aligned}
&\text{(E1)} \quad \frac{d_0}{d_0 + S}\mu_0 + \frac{S}{d_0 + S}\hat{\mu} = \hat{\mu} + \tau(q_l - \hat{\mu}), \\
&\text{(E2)} \quad S\hat{\Sigma} + \nu_0\Sigma_0 + \frac{d_0 S}{d_0 + S}(\mu_0 - \hat{\mu})(\mu_0 - \hat{\mu})^T = (\nu_0 + S - n - 1)\hat{\Sigma}, \\
&\text{(E3)} \quad \frac{1}{d_0 + S}\hat{\Sigma}_B = \tau(1-\tau)\frac{1}{S}\frac{k-2}{k}\hat{\Sigma}_{BL}.
\end{aligned}$$

(E1) can be solved for μ_0 for arbitrary d_0, which yields $\mu_0 = \hat{\mu} + \frac{d_0+S}{d_0}\tau(q_l - \hat{\mu})$. Further, if we set $d_0 = \left(\frac{1}{\tau(1-\tau)}\frac{k}{k-2} - 1\right)S$, then (E3) holds if and only if (E2) holds. Repeating the same arguments as for Theorem 1 proves the claim. □

5 Conclusion

The findings of the last section can be easily summarized to a few core messages:

1. Looking at the traditional portfolio optimization framework, the Black-Litterman is just a special case of the Bayes model. Although this may seem as making the Black-Litterman approach redundant, we still prefer this approach over the more general Bayes approach due to its more intuitive and appealing setup.
2. Considering the robust portfolio optimization framework, both models do not contain the other as special case. In fact, both models can be separated by one feature: while the Bayes model only shortens the classical efficient frontier, the whole shape of the frontier is changed in the Black-Litterman setting. Depending on the purpose in mind, this may clearly favor one approach over the other.
3. If uncertain covariance matrices need to be modelled, the Black-Litterman is no longer suitable. In this case, alternative frameworks like the one of Qian and Gorman, see [25] have to be used.

All in all, we can conclude that none of the two approaches is redundant or superior to the other. The choice of the model has to be made on a case-by-case basis, depending whether it is easier to specify the whole prior distribution or to determine a few forecasts.

References

1. Barry C B (1974) Portfolio analysis under uncertain means, variances and covariances. Journal of Finance 29 2: 512-522
2. Bawa, Brown, Klein (1979) Optimal portfolio choice under uncertainty: A Bayesian approach. In: Estimation Risk and Optimal Portfolio Choice, ed., Bawa, Brown and Klein, North Holland, Amsterdam.
3. Ben-Tal A and Nemirovski (1998) Robust convex optimization. Mathematics of Operations Research 23 4: 769-805
4. Ben-Tal A and Margalit T and Nemirovski A (2000) Robust Modeling of Multi-Stage Portfolio Problems. In: Frenk H and Roos K and Terlaky T and Zhang S (Eds.) High Performance Optimization. Kluwer Academic Publishers
5. Best M J and Grauer R R (1991) On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. The Review of Financial Studies 4 2: 315-342
6. Bevan A and Winkelmann K (1998) Using the Black-Litterman Global Asset Allocation Model: Three Years of Practical Experience. Fixed Income Research, Goldman, Sachs & Co.
7. Black F and Litterman R (1992) Global Portfolio Optimization. Financial Analysts Journal 48 5: 28-43
8. Cornuejols G and Tütüncü R H (2006) Optimization Methods in Finance. Cambridge University Press
9. Drobetz W (2001) How to avoid the pitfalls in portfolio optimization? Putting the Black-Litterman approach at work. Financial Markets and Portfolio Management 15 1: 59-75
10. Efron B and Morris C (1975) Data Analysis Using Stein's Estimator and Its Generalizations. Journal of the American Statistical Association 70: 311-319
11. Fang K-T and Kotz S and Ng K-W (1990) Symmetric Multivariate and Related Distributions. Chapman and Hall, London

12. Frost P A and Savarino J E (1986) An Empirical Bayes Approach to Efficient Portfolio Selection. Journal of Financial and Quantitative Analysis 21 3: 293-305
13. Goldfarb D and Iyengar G (2003) Robust Portfolio Selection Problems. Mathematics of Operations Research 28 1: 1-38
14. He G and Litterman R (1999) The Intuition behind Black-Litterman Model Portfolios. Investment Management Research, Goldman, Sachs & Co., Quantitative Resources Group
15. Idzorek T M (2005) A step-by-step guide to the Black-Litterman model. Working Paper, Available from: http://www.globalriskguard.com/html/virtual_library.html
16. Jorion P (1986) Bayes-Stein Estimation for Portfolio Analysis. Journal of Financial and Quantitative Analysis 21 3
17. Jorion, P (1992) Portfolio Optimization in Practice. Financial Analysts Journal 48 1: 68-74
18. Lutgens F (2004) Robust Portfolio Optimization. PhD Thesis, Maastricht University
19. Markowitz H (1952) Portfolio Selection. Journal of Finance 7 1: 77-91
20. Meucci A (2005) Risk and Asset Allocation. Springer-Verlag, Berlin
21. Meucci A (2005) Robust Bayesian Allocation. SSRN working paper, Available from: http://papers.ssrn.com
22. Press S J (1972) Applied Multivariate Analysis. Holt, Rinehart and Winston, Inc., New York
23. Schöttle K and Werner R (2006) Towards reliable efficient frontiers. Journal of Asset Management 7 2: 128-141
24. Schöttle K and Werner R (2009) Robustness properties of mean-variance portfolios. Optimization, to appear
25. Qian E and Gorman S (2001) Conditional Distribution in Portfolio Theory. Financial Analysts Journal 57 2: 44-51