

## Supplementary material

### Optimal portfolio choice: A minimum expected loss approach

Andrés Ramírez–Hassan · Rosember Guerra–Urzola

Received: date / Accepted: date

**Abstract** The mainstream in finance tackles portfolio selection based on a plug-in approach without consideration of the main objective of the inferential situation. We propose minimum expected loss (MELO) estimators for portfolio selection that explicitly consider the trading rule of interest. The asymptotic properties of our MELO proposal are similar to the plug-in approach. Nevertheless, simulation exercises show that our proposal exhibits better finite sample properties when compared to the competing alternatives, especially when the tangency portfolio is taken as the asset allocation strategy. We have also developed a graphical user interface to help practitioners to use our MELO proposal.

**Keywords** Bayesian estimation · minimum expected loss · portfolio selection

**JEL Classification:** C01; C11; G11.

## 1 Bayesian multivariate regression model

### 1.1 Informative priors

Given the structure of the Treynor–Black model, our point of departure is a multivariate regression model. We then analyze a particular trading strategy from this general framework.<sup>1</sup>

$$\mathbf{R} = \mathbf{X}\mathbf{B} + \mathbf{U},$$

---

Andrés Ramírez–Hassan

Department of Economics, School of Economics and Finance, Universidad EAFIT, Medellín, Colombia.

Tel.: +57-4-2619500

E-mail: aramir21@eafit.edu.co

Rosember Guerra–Urzola

Department of Methodology and Statistics, School of Social and Behavioral Science, Tilburg University.

Tel.: +31-134662965

E-mail: r.i.guerraurzola@uvt.nl

Department of Finance, School of Economics and Finance, Universidad EAFIT, Medellín, Colombia.

Tel.: +57-4-2619500

E-mail: rguerra@eafit.edu.co

<sup>1</sup> Although Bayesian books such as [1,2] present the multivariate regression setting, these books do not present some of the results that we require.

where  $\mathbf{R}$  is a  $T \times N$  matrix,  $\mathbf{X}$  is a  $T \times k$ ,  $\mathbf{B}$  is an  $k \times N$ , and  $\mathbf{U}$  is an  $T \times N$  matrix of disturbance terms with  $\text{vec}(\mathbf{U}) \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_T)$ .

Observe that the Treynor–Black model is a particular case of this framework, such that  $\mathbf{X} = [\mathbf{1} \ \mathbf{r}_M]$

and  $\mathbf{B} = \begin{bmatrix} \boldsymbol{\alpha}' \\ \boldsymbol{\beta}' \end{bmatrix}$ .

**Likelihood function:**

The likelihood function for the model is

$$\mathcal{L}(\mathbf{R} \mid \mathbf{B}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-T/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \mathbf{S} + (\mathbf{B} - \widehat{\mathbf{B}})' \mathbf{X}' \mathbf{X} (\mathbf{B} - \widehat{\mathbf{B}}) \right] \boldsymbol{\Sigma}^{-1} \right\}, \quad (1.1)$$

where  $\mathbf{S} \equiv (\mathbf{R} - \mathbf{X}\widehat{\mathbf{B}})'(\mathbf{R} - \mathbf{X}\widehat{\mathbf{B}})$ ,  $\widehat{\mathbf{B}} \equiv (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R}$  and  $\text{tr}$  is the trace operator.

**Prior distribution:**

The prior distribution of the model parameters is  $\pi(\mathbf{B}, \boldsymbol{\Sigma}) = \pi(\mathbf{B} \mid \boldsymbol{\Sigma})\pi(\boldsymbol{\Sigma})$ , where  $\pi(\mathbf{B} \mid \boldsymbol{\Sigma}) \sim N(\mathbf{B}_0, \boldsymbol{\Sigma} \otimes \mathbf{V}_0)$  and  $\pi(\boldsymbol{\Sigma}) \sim IW(\mathbf{H}_0, \nu_0)$ ; that is:

$$\begin{aligned} \pi(\mathbf{B}, \boldsymbol{\Sigma}) &\propto |\boldsymbol{\Sigma}|^{-k/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ (\mathbf{B} - \mathbf{B}_0)' \mathbf{V}_0^{-1} (\mathbf{B} - \mathbf{B}_0) \right] \boldsymbol{\Sigma}^{-1} \right\} \\ &\times |\boldsymbol{\Sigma}|^{-(\nu_0 + N + 1)/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \mathbf{H}_0 \boldsymbol{\Sigma}^{-1} \right] \right\}, \end{aligned} \quad (1.2)$$

This is the statistical setting for the informative Bayesian approach of the Treynor–Black model.

**Posterior distribution**

Combining the likelihood function (1.1) and the prior distribution (1.2) gives the following posterior distribution:

$$\begin{aligned} \pi(\mathbf{B}, \boldsymbol{\Sigma} \mid \mathbf{R}) &\propto \mathcal{L}(\mathbf{R} \mid \mathbf{B}, \boldsymbol{\Sigma})\pi(\mathbf{B} \mid \boldsymbol{\Sigma})\pi(\boldsymbol{\Sigma}) \\ &\propto |\boldsymbol{\Sigma}|^{-\frac{T+k+\nu_0+N+1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \mathbf{H}_0 + \mathbf{S} + (\mathbf{B} - \mathbf{B}_0)' \mathbf{V}_0^{-1} (\mathbf{B} - \mathbf{B}_0) + (\mathbf{B} - \widehat{\mathbf{B}})' \mathbf{X}' \mathbf{X} (\mathbf{B} - \widehat{\mathbf{B}}) \right] \right\}. \end{aligned}$$

Completing the squares on  $\mathbf{B}$  and collecting the remaining terms in the bracket yields

$$\begin{aligned} \mathbf{H}_0 + \mathbf{S} + (\mathbf{B} - \mathbf{B}_0)' \mathbf{V}_0^{-1} (\mathbf{B} - \mathbf{B}_0) + (\mathbf{B} - \widehat{\mathbf{B}})' \mathbf{X}' \mathbf{X} (\mathbf{B} - \widehat{\mathbf{B}}) \\ = (\mathbf{B} - \widetilde{\mathbf{B}})' \widetilde{\mathbf{V}}^{-1} (\mathbf{B} - \widetilde{\mathbf{B}}) + \widetilde{\mathbf{H}}, \end{aligned}$$

where

$$\begin{aligned} \widetilde{\mathbf{B}} &\equiv (\mathbf{V}_0^{-1} + \mathbf{X}'\mathbf{X})^{-1}(\mathbf{V}_0^{-1}\mathbf{B}_0 + \mathbf{X}'\mathbf{R}), \\ \widetilde{\mathbf{V}} &\equiv (\mathbf{V}_0^{-1} + \mathbf{X}'\mathbf{X})^{-1}, \\ \widetilde{\mathbf{H}} &\equiv \mathbf{H}_0 + \mathbf{S} + \mathbf{B}_0' \mathbf{V}_0^{-1} \mathbf{B}_0 + \widehat{\mathbf{B}}' \mathbf{X}' \mathbf{X} \widehat{\mathbf{B}} - \widetilde{\mathbf{B}}' \widetilde{\mathbf{V}}^{-1} \widetilde{\mathbf{B}}. \end{aligned}$$

Thus, the posterior distribution can be written as

$$\begin{aligned} \pi(\mathbf{B}, \boldsymbol{\Sigma} \mid \mathbf{R}) &\propto |\boldsymbol{\Sigma}|^{-k/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ (\mathbf{B} - \widetilde{\mathbf{B}})' \widetilde{\mathbf{V}}^{-1} (\mathbf{B} - \widetilde{\mathbf{B}}) \right] \boldsymbol{\Sigma}^{-1} \right\} \\ &\times |\boldsymbol{\Sigma}|^{-\frac{T+\nu_0+N+1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \widetilde{\mathbf{H}} \boldsymbol{\Sigma}^{-1} \right] \right\}. \end{aligned} \quad (1.3)$$

That is  $\pi(\mathbf{B}, \boldsymbol{\Sigma} | \mathbf{R}) \propto \pi(\mathbf{B} | \boldsymbol{\Sigma}, \mathbf{R})\pi(\boldsymbol{\Sigma} | \mathbf{R})$  where

$$\begin{aligned}\pi(\mathbf{B} | \boldsymbol{\Sigma}, \mathbf{R}) &\sim N(\tilde{\mathbf{B}}, \boldsymbol{\Sigma} \otimes \tilde{\mathbf{V}}), \\ \pi(\boldsymbol{\Sigma} | \mathbf{R}) &\sim IW(\tilde{\mathbf{H}}, \tilde{\nu}),\end{aligned}$$

where  $\tilde{\nu} = T + \nu_0$ .

### Predictive distribution

The point of departure of the predictive density for the next  $\kappa$  times periods is

$$\mathbf{R}_{T+\kappa} = \mathbf{Z}\mathbf{B} + \mathbf{V},$$

where  $\mathbf{R}_{T+\kappa}$  is a  $\kappa \times N$  matrix,  $\mathbf{Z}$  is a  $\kappa \times k$  matrix. Then, the predictive pdf for  $\mathbf{R}_{T+\kappa}$  is given by

$$p(\mathbf{R}_{T+\kappa} | \mathbf{R}, \mathbf{X}, \mathbf{Z}) \propto \int \int p(\mathbf{B}, \boldsymbol{\Sigma} | \mathbf{R}, \mathbf{X})p(\mathbf{R}_{T+\kappa} | \mathbf{Z}, \mathbf{B}, \boldsymbol{\Sigma})d\mathbf{B}d\boldsymbol{\Sigma}, \quad (1.4)$$

where

$$p(\mathbf{R}_{T+\kappa} | \mathbf{Z}, \mathbf{B}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-\kappa/2} \exp \left\{ -\frac{1}{2} \text{tr} [(\mathbf{R}_{T+\kappa} - \mathbf{Z}\mathbf{B})' (\mathbf{R}_{T+\kappa} - \mathbf{Z}\mathbf{B}) \boldsymbol{\Sigma}^{-1}] \right\}, \quad (1.5)$$

combining (1.3) and (1.5) we get,

$$|\boldsymbol{\Sigma}|^{-\frac{k+\tilde{\nu}+N+\kappa+1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} [\mathbf{A}\boldsymbol{\Sigma}^{-1}] \right\},$$

where  $\mathbf{A} = (\mathbf{B} - \tilde{\mathbf{B}})' \tilde{\mathbf{V}}^{-1} (\mathbf{B} - \tilde{\mathbf{B}}) + \tilde{\mathbf{H}} + (\mathbf{R}_{T+\kappa} - \mathbf{Z}\mathbf{B})' (\mathbf{R}_{T+\kappa} - \mathbf{Z}\mathbf{B})$ . We use the inverse wishart properties to integrate (1.4) with respect to  $\boldsymbol{\Sigma}$ ,

$$|\mathbf{A}|^{-\frac{k+\tilde{\nu}+\kappa}{2}} = \left| \tilde{\mathbf{H}} + (\mathbf{B} - \tilde{\mathbf{B}})' \tilde{\mathbf{V}}^{-1} (\mathbf{B} - \tilde{\mathbf{B}}) + (\mathbf{R}_{T+\kappa} - \mathbf{Z}\mathbf{B})' (\mathbf{R}_{T+\kappa} - \mathbf{Z}\mathbf{B}) \right|^{-\frac{k+\tilde{\nu}+\kappa}{2}}.$$

To integrate with respect to  $\mathbf{B}$ , we complete the square on  $\mathbf{B}$ ,

$$\begin{aligned}A &= \tilde{\mathbf{H}} + \mathbf{B}' \tilde{\mathbf{V}}^{-1} \mathbf{B} - 2\mathbf{B}' \tilde{\mathbf{V}}^{-1} \tilde{\mathbf{B}} + \tilde{\mathbf{B}}' \tilde{\mathbf{V}}^{-1} \tilde{\mathbf{B}} \\ &\quad + \mathbf{R}'_{T+\kappa} \mathbf{R}_{T+\kappa} - 2\mathbf{B}' \mathbf{Z}' \mathbf{R}_{T+\kappa} + \mathbf{B}' \mathbf{Z}' \mathbf{Z} \mathbf{B} \\ &= \tilde{\mathbf{H}} + \tilde{\mathbf{B}}' \tilde{\mathbf{V}}^{-1} \tilde{\mathbf{B}} + \mathbf{R}'_{T+\kappa} \mathbf{R}_{T+\kappa} - \mathbf{B}^* \mathbf{M} \mathbf{B}^* \\ &\quad + \mathbf{B}' \mathbf{M} \mathbf{B} - 2\mathbf{B}' \mathbf{M} \mathbf{B}^* + \mathbf{B}^{*'} \mathbf{M} \mathbf{B}^* \\ &= \tilde{\mathbf{H}} + \tilde{\mathbf{B}}' \tilde{\mathbf{V}}^{-1} \tilde{\mathbf{B}} + \mathbf{R}'_{T+\kappa} \mathbf{R}_{T+\kappa} - \mathbf{B}^{*'} \mathbf{M} \mathbf{B}^* + (\mathbf{B} - \mathbf{B}^*)' \mathbf{M} (\mathbf{B} - \mathbf{B}^*),\end{aligned}$$

where  $\mathbf{M} = \tilde{\mathbf{V}}^{-1} + \mathbf{Z}' \mathbf{Z}$  and  $\mathbf{B}^* = \mathbf{M}^{-1} (\tilde{\mathbf{B}}' \tilde{\mathbf{V}}^{-1} + \mathbf{R}'_{T+\kappa} \mathbf{Z})$ .

Now, using properties of the generalized multivariate Student- $t$  pdf, we can perform the integration with respect to the parameter  $\mathbf{B}$ , which yields,

$$p(\mathbf{R}_{T+\kappa} | \mathbf{R}, \mathbf{X}, \mathbf{Z}) \propto \left| \tilde{\mathbf{H}} + \tilde{\mathbf{B}}' \tilde{\mathbf{V}}^{-1} \tilde{\mathbf{B}} + \mathbf{R}'_{T+\kappa} \mathbf{R}_{T+\kappa} - \mathbf{B}^{*'} \mathbf{M} \mathbf{B}^* \right|^{-\frac{\tilde{\nu}+\kappa}{2}}. \quad (1.6)$$

To simplify this expression we complete the square on  $\mathbf{R}_{T+\kappa}$  as follows,

$$\begin{aligned}
\tilde{\mathbf{B}}'\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{B}} + \mathbf{R}'_{T+\kappa}\mathbf{R}_{T+\kappa} - \mathbf{B}^{*\prime}\mathbf{M}\mathbf{B}^* &= \tilde{\mathbf{B}}'\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{B}} + \mathbf{R}'_{T+\kappa}\mathbf{R}_{T+\kappa} - \tilde{\mathbf{B}}'\tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{B}} \\
&\quad - \tilde{\mathbf{B}}'\tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\mathbf{Z}'\mathbf{R}_{T+\kappa} - \mathbf{R}_{T+\kappa}\mathbf{Z}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{B}} - \mathbf{R}'_{T+\kappa}\mathbf{Z}\mathbf{Z}'\mathbf{R}_{T+\kappa} \\
&= \tilde{\mathbf{B}}'(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1})\tilde{\mathbf{B}} - \mathbf{B}^{**\prime}\mathbf{C}\mathbf{B}^{**} \\
&\quad + (\mathbf{R}_{T+\kappa} - \mathbf{B}^{**})'\mathbf{C}(\mathbf{R}_{T+\kappa} - \mathbf{B}^{**}), \tag{1.7}
\end{aligned}$$

where  $\mathbf{C} = \mathbf{I} - \mathbf{Z}\mathbf{M}^{-1}\mathbf{Z}'$ ,  $\mathbf{C}^{-1} = \mathbf{I} + \mathbf{Z}\tilde{\mathbf{V}}\mathbf{Z}'$  and  $\mathbf{B}^{**} = \mathbf{C}^{-1}\mathbf{Z}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{B}}$ . So, let us see what  $\mathbf{C}^{-1}\mathbf{Z}\mathbf{M}^{-1}$  is

$$\begin{aligned}
\mathbf{C}^{-1}\mathbf{Z}\mathbf{M}^{-1} &= [\mathbf{I} - \mathbf{Z}\mathbf{M}^{-1}\mathbf{Z}']^{-1}\mathbf{Z}\mathbf{M}^{-1} \\
&= \mathbf{Z}[\mathbf{I} + \tilde{\mathbf{V}}\mathbf{Z}'\mathbf{Z}]\mathbf{M}^{-1} \\
&= \mathbf{Z}\tilde{\mathbf{V}}(\tilde{\mathbf{V}}^{-1} + \mathbf{Z}'\mathbf{Z})\mathbf{M}^{-1} \\
&= \mathbf{Z}\tilde{\mathbf{V}}. \tag{1.8}
\end{aligned}$$

Then,  $\mathbf{B}^{**} = \mathbf{Z}\tilde{\mathbf{B}}$  and

$$\begin{aligned}
\tilde{\mathbf{B}}'(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1})\tilde{\mathbf{B}} - \mathbf{B}^{**\prime}\mathbf{C}\mathbf{B}^{**} &= \tilde{\mathbf{B}}'(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\mathbf{Z}'\mathbf{C}^{-1}\mathbf{Z}\mathbf{M}^{-1}\tilde{\mathbf{V}}^{-1})\tilde{\mathbf{B}} \\
&= \tilde{\mathbf{B}}(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}[\mathbf{I} + \mathbf{Z}'\mathbf{Z}\tilde{\mathbf{V}}]\tilde{\mathbf{V}}^{-1})\tilde{\mathbf{B}} \\
&= \tilde{\mathbf{B}}(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}[\tilde{\mathbf{V}}^{-1} + \mathbf{Z}'\mathbf{Z}])\tilde{\mathbf{B}} \\
&= \tilde{\mathbf{B}}(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{M}^{-1}\mathbf{M})\tilde{\mathbf{B}} \\
&= \tilde{\mathbf{B}}(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1})\tilde{\mathbf{B}} = \mathbf{0}. \tag{1.9}
\end{aligned}$$

Finally, by substituting (1.8) and (1.9) in (1.7) we get,

$$\tilde{\mathbf{H}} + \tilde{\mathbf{B}}'\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{B}} + \mathbf{R}'_{T+\kappa}\mathbf{R}_{T+\kappa} - \mathbf{B}^{*\prime}\mathbf{M}\mathbf{B}^* = \tilde{\mathbf{H}} + (\mathbf{R}_{T+\kappa} - \mathbf{Z}\tilde{\mathbf{B}})'(\mathbf{I} - \mathbf{Z}\mathbf{M}^{-1}\mathbf{Z}')(\mathbf{R}_{T+\kappa} - \mathbf{Z}\tilde{\mathbf{B}}).$$

We can write the predictive pdf (1.6) as

$$p(\mathbf{R}_{T+\kappa} \mid \mathbf{R}, \mathbf{X}, \mathbf{Z}) \propto \left| \tilde{\mathbf{H}} + (\mathbf{R}_{T+\kappa} - \mathbf{Z}\tilde{\mathbf{B}})'(\mathbf{I} - \mathbf{Z}\mathbf{M}^{-1}\mathbf{Z}')(\mathbf{R}_{T+\kappa} - \mathbf{Z}\tilde{\mathbf{B}}) \right|^{-\frac{\tilde{\nu}+\kappa}{2}}. \tag{1.10}$$

The matrix of future observations has a pdf in the generalized multivariate Student- $t$  distribution with  $\tilde{\nu} + \kappa - N$  degrees of freedom.

This distribution was introduced by [3, 4] and is known as matrix  $t$  pdf. [1] present some properties of (1.10) but in our work we only use one of these properties.

Let

$$\mathbf{C} = \mathbf{I} - \mathbf{Z}\mathbf{M}^{-1}\mathbf{Z}' = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix},$$

and

$$\mathbf{C}_{22 \cdot 1} = \mathbf{C}_{22} - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}, \tag{1.11}$$

where  $\mathbf{C}_{11}$  is a block matrix of dimension  $p_1 \times p_1$ ,  $\mathbf{C}_{22}$  is of dimension  $p_2 \times p_2$ , and  $p_1 + p_2 = \kappa$ . Then, the marginal pdf for  $\mathbf{R}_{T+\kappa, p_2}$  is

$$p(\mathbf{R}_{T+\kappa; p_2} | \mathbf{R}, \mathbf{X}, \mathbf{Z}) \propto \left| \widetilde{\mathbf{H}} + (\mathbf{R}_{T+\kappa, p_2} - \mathbf{Z}_{p_2} \widetilde{\mathbf{B}})' (\mathbf{I} - \mathbf{Z}_{p_2} \mathbf{M}_{p_2}^{-1} \mathbf{Z}_{p_2}') (\mathbf{R}_{T+\kappa; p_2} - \mathbf{Z}_{p_2} \widetilde{\mathbf{B}}) \right|^{-\frac{\bar{\nu} + \kappa}{2}}. \quad (1.12)$$

In the last notation, the subindex  $\mathbf{R}_{T+\kappa, p_2}$  means the last  $p_2$  rows of the  $\mathbf{R}_{T+\kappa}$  matrix. In the particular case of Treynor–Black approach with  $p_2 = \kappa$ , the mean and variance of the predictive distribution are presented in equations 2.10 and 2.11 in our paper, respectively.

Observe that to get informative Bayesian estimates for  $\boldsymbol{\mu}_{T+\kappa}$  and  $\boldsymbol{\Sigma}_{T+\kappa}$  in the portfolio optimization problem, we can define the multivariate regression setting as

$$\mathbf{R} = \mathbf{1}\boldsymbol{\mu}' + \mathbf{U},$$

where  $\mathbf{R}$  is a  $T \times N$  matrix,  $\boldsymbol{\mu}$  a  $N \times 1$  vector,  $\mathbf{U}$  a  $T \times N$  matrix of disturbance terms. We use the following informative priors for the excess of return and the covariance matrix:

$$\begin{aligned} \boldsymbol{\mu} | \boldsymbol{\Sigma} &\sim N\left(\boldsymbol{\eta}, \frac{1}{\tau} \boldsymbol{\Sigma}\right), \\ \boldsymbol{\Sigma} &\sim IW(\boldsymbol{\Omega}, \nu). \end{aligned}$$

The likelihood function is

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma} | \mathbf{R}) = |\boldsymbol{\Sigma}|^{-T/2} \exp\left\{-\frac{1}{2} \text{tr}[\mathbf{S} + T(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})'] \boldsymbol{\Sigma}^{-1}\right\},$$

where  $\mathbf{S} = (\mathbf{R} - \mathbf{1}\hat{\boldsymbol{\mu}})'(\mathbf{R} - \mathbf{1}\hat{\boldsymbol{\mu}}')$  and  $\hat{\boldsymbol{\mu}}$  is the maximum likelihood estimator. The Bayes' rule implies

$$\begin{aligned} \pi(\boldsymbol{\mu}, \boldsymbol{\Sigma} | \mathbf{R}) &= \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma} | \mathbf{R}) \pi(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma} | \mathbf{R}) \pi(\boldsymbol{\mu} | \boldsymbol{\Sigma}) \pi(\boldsymbol{\Sigma}) \\ &= |\boldsymbol{\Sigma}|^{-\frac{T+\nu+N+2}{2}} \exp\left\{-\frac{1}{2} \text{tr}[\boldsymbol{\Omega} + \mathbf{S} + T(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})' + \tau(\boldsymbol{\mu} - \boldsymbol{\eta})(\boldsymbol{\mu} - \boldsymbol{\eta})'] \boldsymbol{\Sigma}^{-1}\right\}. \end{aligned} \quad (1.13)$$

Expanding the last two terms in the exponential and completing squares in  $\boldsymbol{\mu}$ ,

$$\begin{aligned} T(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})' + \tau(\boldsymbol{\mu} - \boldsymbol{\eta})(\boldsymbol{\mu} - \boldsymbol{\eta})' &= T[\boldsymbol{\mu}\boldsymbol{\mu}' - \boldsymbol{\mu}\hat{\boldsymbol{\mu}}' - \hat{\boldsymbol{\mu}}\boldsymbol{\mu}' + \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}'] + \tau[\boldsymbol{\mu}\boldsymbol{\mu}' - \boldsymbol{\mu}\boldsymbol{\eta}' - \boldsymbol{\eta}\boldsymbol{\mu}' + \boldsymbol{\eta}\boldsymbol{\eta}'] \\ &= (T + \tau)\boldsymbol{\mu}\boldsymbol{\mu}' - 2(T\boldsymbol{\mu}\hat{\boldsymbol{\mu}}' + \tau\boldsymbol{\eta}\boldsymbol{\eta}') + T\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}' + \tau\boldsymbol{\eta}\boldsymbol{\eta}' \\ &= (T + \tau)(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})' + T\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}' + \tau\boldsymbol{\eta}\boldsymbol{\eta}' - (T + \tau)\tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\mu}}' \\ &= (T + \tau)(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})' + T\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}' + \tau\boldsymbol{\eta}\boldsymbol{\eta}' - \frac{1}{T + \tau}(T\hat{\boldsymbol{\mu}} + \tau\boldsymbol{\eta})(T\hat{\boldsymbol{\mu}} + \tau\boldsymbol{\eta}) \\ &= (T + \tau)(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})' + \frac{T\tau}{T + \tau}(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})', \end{aligned}$$

where  $\tilde{\boldsymbol{\mu}} = \frac{T}{T+\tau}\hat{\boldsymbol{\mu}} + \frac{\tau}{T+\tau}\boldsymbol{\eta}$ , a weighted average between the prior mean and the sample mean of the returns. Then, equation (1.13) is:

$$\begin{aligned} & |\boldsymbol{\Sigma}|^{-\frac{T+\nu+N+2}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \boldsymbol{\Omega} + (T-1)\widehat{\boldsymbol{\Sigma}} + (T+\tau)(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})' + \frac{T\tau}{T+\tau}(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})' \right] \boldsymbol{\Sigma}^{-1} \right\} \\ &= |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[ (T+\tau)(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})(\boldsymbol{\mu} - \tilde{\boldsymbol{\mu}})' \right] \boldsymbol{\Sigma}^{-1} \right\} \\ &\times |\boldsymbol{\Sigma}|^{-\frac{T+\nu+N+1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \boldsymbol{\Omega} + (T-1)\widehat{\boldsymbol{\Sigma}} + \frac{T\tau}{T+\tau}(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})' \right] \boldsymbol{\Sigma}^{-1} \right\}. \end{aligned}$$

Thus

$$\pi(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{R}) = \pi(\boldsymbol{\mu} \mid \boldsymbol{\Sigma}, \mathbf{R}) \times \pi(\boldsymbol{\Sigma} \mid \mathbf{R}),$$

where

$$\begin{aligned} \boldsymbol{\mu} \mid \boldsymbol{\Sigma}, \mathbf{R} &\sim N \left( \tilde{\boldsymbol{\mu}}, \frac{1}{T+\tau} \boldsymbol{\Sigma} \right), \\ \boldsymbol{\Sigma} \mid \mathbf{R} &\sim IW \left( \tilde{\boldsymbol{\Omega}}, \tilde{\nu} \right), \end{aligned}$$

where  $\tilde{\boldsymbol{\Omega}} = \boldsymbol{\Omega} + (T-1)\widehat{\boldsymbol{\Sigma}} + \frac{T\tau}{T+\tau}(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})(\hat{\boldsymbol{\mu}} - \boldsymbol{\eta})'$  and  $\tilde{\nu} = \nu + T$ .

Then, setting  $\tilde{\mathbf{H}} = \tilde{\boldsymbol{\Omega}}$ ,  $\tilde{\nu} = T + \nu$ ,  $\tilde{\mathbf{B}} = \tilde{\boldsymbol{\mu}}$ ,  $\mathbf{Z} = \mathbf{1}_\kappa$ , and  $\tilde{\mathbf{V}} = \frac{1}{T+\tau}$  in the equation (1.10), we have the following specification for the predictive distribution

$$p(\mathbf{R}_{T+\kappa} \mid \mathbf{R}) \propto \left| \tilde{\boldsymbol{\Omega}} + (\mathbf{R}_{T+\kappa} - \mathbf{1}_\kappa \tilde{\boldsymbol{\mu}})' \left( \mathbf{I}_\kappa - \frac{\mathbf{1}_\kappa \mathbf{1}_\kappa'}{T+\tau+\kappa} \right) (\mathbf{R}_{T+\kappa} - \mathbf{1}_\kappa \tilde{\boldsymbol{\mu}}) \right|^{-\frac{T+\nu+\kappa-1}{2}}$$

This is a multivariate Student's t-distribution with  $T + \nu + \kappa - N$  degrees of freedom.

The following step is to use the marginal distribution described in equation (1.12) with  $p_1 = \kappa - 1$ .

Thus, the marginal pdf for the  $\kappa$ -row of  $\mathbf{R}_{T+\kappa}$  is given by

$$p(\mathbf{R}_{T+\kappa;\kappa} \mid \mathbf{R}) \propto \left| 1 + \mathbf{C}_{22 \cdot 1} (\mathbf{R}_{T+\kappa;\kappa} - \tilde{\boldsymbol{\mu}})' \tilde{\boldsymbol{\Omega}}^{-1} (\mathbf{R}_{T+\kappa;\kappa} - \tilde{\boldsymbol{\mu}}) \right|^{-\frac{T+\nu+\kappa-1}{2}},$$

where

$$\mathbf{C}_{22 \cdot 1} = \frac{(T+\tau)(T+\tau+\kappa)(T+\tau+\kappa-1) - (\kappa-1)(T+\tau-\kappa+1)}{(T+\tau)(T+\tau+\kappa)^2}.$$

Thus, the mean and variance of the distribution are presented in equation 2.8 and 2.9 in our paper, respectively.

## 1.2 Non-informative priors

The non-informative case uses a Jeffrey's type prior

$$\pi(\mathbf{B}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(N+1)/2}.$$

Then, the conditional posterior distributions are

$$\begin{aligned} \pi(\mathbf{B} \mid \boldsymbol{\Sigma}, \mathbf{R}) &\sim N(\widehat{\mathbf{B}}, \boldsymbol{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1}), \\ \pi(\boldsymbol{\Sigma} \mid \mathbf{R}) &\sim IW(\mathbf{S}, T-2). \end{aligned}$$

By setting  $\widetilde{\mathbf{H}} = \mathbf{S}$ ,  $\tilde{\nu} = T - 2$ ,  $\widetilde{\mathbf{B}} = \widehat{\mathbf{B}}$ , and  $\widetilde{\mathbf{V}} = (\mathbf{X}'\mathbf{X})^{-1}$  in the equation (1.10) we have the following specification for the predictive distribution

$$p(\mathbf{R}_{T+\kappa} | \mathbf{R}) \propto \left| \mathbf{S} + (\mathbf{R}_{T+\kappa} - \mathbf{Z}\widehat{\mathbf{B}})' \mathbf{C} (\mathbf{R}_{T+\kappa} - \mathbf{Z}\widehat{\mathbf{B}}) \right|^{-\frac{T+\kappa-2}{2}},$$

where  $\mathbf{C} = \mathbf{I} - \mathbf{Z}\mathbf{M}^{-1}\mathbf{Z}'$ .

This is a multivariate student-t distribution with  $T + \kappa - N - 2$  degrees of freedom.

The following step is using the marginal distribution described in equation (1.12) with  $p_1 = \kappa - 1$ .

Thus, the marginal pdf for the  $\kappa$ -row of  $\mathbf{R}_{T+\kappa}$  is given by

$$p(\mathbf{R}_{T+\kappa;\kappa} | \mathbf{R}) \propto \left| 1 + \mathbf{C}_{22\cdot 1} (\mathbf{R}_{T+\kappa;\kappa} - \mathbf{Z}_\kappa \widehat{\boldsymbol{\mu}})' \mathbf{S}^{-1} (\mathbf{R}_{T+\kappa;\kappa} - \mathbf{Z}_\kappa \widehat{\boldsymbol{\mu}}) \right|^{-\frac{T+\nu+\kappa-1}{2}},$$

where  $\mathbf{C}_{22\cdot 1}$  defined as in equation (1.11). The mean and variance of the distribution are presented in equations 2.6 and 2.7 in our paper, respectively.

Observe that to get non-informative Bayesian estimates for  $\boldsymbol{\mu}_{T+\kappa}$  and  $\boldsymbol{\Sigma}_{T+\kappa}$ , we can again define the multivariate regression setting as

$$\mathbf{R} = \mathbf{1}\boldsymbol{\mu}' + \mathbf{U},$$

where  $\boldsymbol{\mu}' = \mathbf{B}$  a  $1 \times N$  vector,  $\mathbf{1} = \mathbf{X}$ , and  $k = 1$ . The likelihood function is given by

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-T/2} \exp \left\{ -\frac{1}{2} \text{tr} [S + T(\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})'(\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})] \boldsymbol{\Sigma}^{-1} \right\},$$

where  $S \equiv (\mathbf{R} - \mathbf{1}\widehat{\boldsymbol{\mu}})'(\mathbf{R} - \mathbf{1}\widehat{\boldsymbol{\mu}})$  and  $\widehat{\boldsymbol{\mu}} \equiv \frac{1}{T} \sum_{t=1}^T \mathbf{R}_t$ .

Now, if we consider the non-informative prior,  $\pi(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Sigma}|^{-(N+1)/2}$ , then the joint posterior distribution is

$$\pi(\boldsymbol{\mu}, \boldsymbol{\Sigma} | \mathbf{R}) \propto |\boldsymbol{\Sigma}|^{-(N+T+1)/2} \exp \left\{ -\frac{1}{2} \text{tr} [S + T(\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})'(\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})] \boldsymbol{\Sigma}^{-1} \right\}.$$

Thus, the marginal conditional distributions are

$$\begin{aligned} \pi(\boldsymbol{\mu} | \boldsymbol{\Sigma}, \mathbf{R}) &\sim N \left( \widehat{\boldsymbol{\mu}}, \frac{1}{T} \boldsymbol{\Sigma} \right), \\ \pi(\boldsymbol{\Sigma} | \mathbf{R}) &\sim IW(S, T - 1). \end{aligned}$$

Then, by setting  $\widetilde{\mathbf{H}} = \mathbf{S}$ ,  $\tilde{\nu} = T - 1$ ,  $\widetilde{\mathbf{B}} = \widehat{\boldsymbol{\mu}}$ ,  $\mathbf{Z} = \mathbf{1}_\kappa$ , and  $\widetilde{\mathbf{V}} = \frac{1}{T}$  in the equation (1.10), we have the following specification for the predictive distribution

$$p(\mathbf{R}_{T+\kappa} | \mathbf{R}) \propto \left| \mathbf{S} + (\mathbf{R}_{T+\kappa} - \mathbf{1}\widehat{\boldsymbol{\mu}})' \left( \mathbf{I}_\kappa - \frac{\mathbf{1}_\kappa \mathbf{1}_\kappa'}{T + \kappa} \right) (\mathbf{R}_{T+\kappa} - \mathbf{1}\widehat{\boldsymbol{\mu}}) \right|^{-\frac{T+\kappa-1}{2}}.$$

This is a multivariate Student's t-distribution with  $T + \kappa - N - 1$  degrees of freedom.

The following step is to use the marginal distribution described in equation (1.12) with  $p_1 = \kappa - 1$ .

Thus, the marginal pdf for the  $\kappa$ -row of  $\mathbf{R}_{T+\kappa}$  is given by

$$p(\mathbf{R}_{T+\kappa;\kappa} | \mathbf{R}) \propto \left| 1 + \mathbf{C}_{22\cdot 1} (\mathbf{R}_{T+\kappa;\kappa} - \widehat{\boldsymbol{\mu}})' \mathbf{S}^{-1} (\mathbf{R}_{T+\kappa;\kappa} - \widehat{\boldsymbol{\mu}}) \right|^{-\frac{T+\kappa-1}{2}},$$

where

$$\mathbf{C}_{22\cdot 1} = \frac{T(T + \kappa)(T + \kappa - 1) - (\kappa - 1)(T - \kappa + 1)}{T(T + \kappa)^2}.$$

whose mean and variance are equations 2.4 and 2.5 in our paper, respectively.

## 2 Minimum expected loss estimator for trading strategies

### 2.1 Assumptions

Let  $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_T$  be iid each with density  $f(\mathbf{R}|\boldsymbol{\theta})$  with respect to a  $\sigma$ -finite measure  $\mu$ , where  $\boldsymbol{\theta}$  is real-valued, and suppose the following regularity conditions hold.

#### A. Likelihood

1. The parameter space  $\Theta$  is an open subset of  $\mathbb{R}^L$ .
2. The set  $A = \{\mathbf{R} : f(\mathbf{R}|\boldsymbol{\theta}) > 0\}$  is independent of  $\boldsymbol{\theta}$ .
3. For every  $\mathbf{R} \in A$ , the density  $f(\mathbf{R}|\boldsymbol{\theta})$  is twice differentiable with respect to  $\boldsymbol{\theta}$ , and the second derivative is continuous in  $\boldsymbol{\theta}$ .
4. The Fisher information  $I(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}} \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \log f(\mathbf{R}|\boldsymbol{\theta}) \frac{\partial}{\partial \boldsymbol{\theta}'} \log f(\mathbf{R}|\boldsymbol{\theta}) \right]$  satisfies  $0 < [I(\boldsymbol{\theta})_{ij}] < \infty, i, j = 1, 2, \dots, L$ , where  $[A_{ij}]$  denotes element  $ij$  of matrix  $\mathbf{A}$ .
5. The integral  $\int f(\mathbf{R}|\boldsymbol{\theta}) d\mu(\mathbf{R})$  can be twice differentiated with respect to  $\boldsymbol{\theta}$  under the integral sign. This ensures that for all  $\boldsymbol{\theta} \in \Theta$ ,  $E \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \log f(\mathbf{R}|\boldsymbol{\theta}) \right] = \mathbf{0}$  and  $E \left[ -\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \log f(\mathbf{R}|\boldsymbol{\theta}) \right] = I(\boldsymbol{\theta})$ .
6. For any given  $\boldsymbol{\theta}_0 \in \Theta$ , there exists a positive number  $c$  and a function  $M(\mathbf{R})$  (both of which may depend on  $\boldsymbol{\theta}_0$ ) such that  $|\partial^2 \log f(\mathbf{R}|\boldsymbol{\theta}) / \partial \theta_i \partial \theta_j| \leq M(\mathbf{R})$  for all  $\mathbf{R} \in A$ ,  $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < c$ , where  $\|\cdot\|$  is the Euclidean norm, and  $\mathbb{E}_{\boldsymbol{\theta}_0} M(\mathbf{R}) < \infty$ .

Under these assumptions, it is well-known that the maximum likelihood estimator satisfies  $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, I(\boldsymbol{\theta}_0)^{-1})$ , that is,  $\hat{\boldsymbol{\theta}}$  is consistent for the true value  $\boldsymbol{\theta}_0$ , and asymptotically efficient ([5]). Then,  $\frac{1}{T}[R_T(\boldsymbol{\theta})_{ij}] \xrightarrow{p} \mathbf{0}$  in the second order Taylor series expansion

$$l(\boldsymbol{\theta}) = l(\boldsymbol{\theta}_0) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \frac{\partial l}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}_0} + \frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)' [TI(\boldsymbol{\theta}_0) + R_T(\boldsymbol{\theta})] (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \quad (2.1)$$

where  $l(\boldsymbol{\theta}) = \log(f(\mathbf{R}|\boldsymbol{\theta}))$  is the log likelihood,  $\mathbf{R} = [R_1, R_2, \dots, R_T]$ . However, Bayesian estimators involve an integral over the whole range of  $\boldsymbol{\theta}$  values, this necessitates the following assumption.

#### B. Taylor series expansion

1. Given any  $\epsilon > 0$ , there exist  $\delta > 0$  such that in the expansion 2.1,

$$\lim_{T \rightarrow \infty} P \left( \sup \left\{ \left| \frac{1}{T} [R_T(\boldsymbol{\theta})_{ij}] \right| : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \delta \right\} \geq \epsilon \right) = 0.$$

#### C. Log likelihood bounded contribution

1. For any  $\delta > 0$ , there exist  $\epsilon > 0$  such that

$$\lim_{T \rightarrow \infty} P \left( \sup \left\{ \frac{1}{T} [l(\boldsymbol{\theta}) - l(\boldsymbol{\theta}_0)] : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \geq \delta \right\} \leq -\epsilon \right) = 1.$$

This assumption implies that the log likelihood contribution of  $\boldsymbol{\theta} \notin \mathcal{B}_\delta(\boldsymbol{\theta}_0)$ , where  $\mathcal{B}_\delta(\boldsymbol{\theta}_0)$  is an open ball centered at  $\boldsymbol{\theta}_0$  with radius  $\delta$ , is negligible as  $T \rightarrow \infty$ .



#### D. Prior density

1. The prior density  $\pi(\boldsymbol{\theta})$  is continuous and positive for all  $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ .
2. The expectation and second moment of  $\boldsymbol{\theta}$  under  $\pi$  exists; that is,  $\int \|\boldsymbol{\theta}\|^2 \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} < \infty$ .

The first assumption implies  $\pi(\boldsymbol{\theta}_0) > 0$ , so  $\boldsymbol{\theta}_0$  is not a priori excluded. The second assumption is required to proof that the minimum expected loss estimator is consistent and asymptotically efficient. We assume that there is a proper prior density function. However, result 2.2 can be extended to the case  $\int \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} = \infty$  when there is  $n_0$ , such that the posterior density taking information up to this point is a proper density with probability 1, satisfying assumptions **D**.

#### E. Objective function

1.  $g_i(\boldsymbol{\theta}) = \frac{l_i(\boldsymbol{\theta})}{m(\boldsymbol{\theta})} : \mathbb{R}^L \rightarrow \mathbb{R}$  is a function with finite and nonzero first order derivative at  $\boldsymbol{\theta}_0$ ,  $l_i(\boldsymbol{\theta})$  and  $m(\boldsymbol{\theta}) \neq 0$  are polynomial functions in  $\boldsymbol{\theta}$ , and  $g_i(\boldsymbol{\theta}_0) \neq 0$ , such that  $g_i(\boldsymbol{\theta}) = g_i(\boldsymbol{\theta}_0) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)'[\nabla g_i(\boldsymbol{\theta}_0) + W_T(\boldsymbol{\theta})]$ , and  $\sup \{\|W_T(\boldsymbol{\theta})\| : \boldsymbol{\theta} \in \boldsymbol{\Theta}\} < c_1 < \infty, T \rightarrow \infty$ .

#### F. Weighting functions

1.  $h(\boldsymbol{\theta}) : \mathbb{R}^L \rightarrow \mathbb{R}^{++}$  is a function with finite and nonzero first order derivative at  $\boldsymbol{\theta}_0$ , continuous first order derivative, and  $h(\boldsymbol{\theta}_0) \neq 0$ , such that  $h(\boldsymbol{\theta}) = h(\boldsymbol{\theta}_0) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)'[\nabla h(\boldsymbol{\theta}_0) + V_T(\boldsymbol{\theta})]$ , and  $\sup \{\|V_T(\boldsymbol{\theta})\| : \boldsymbol{\theta} \in \boldsymbol{\Theta}\} < c_2 < \infty, T \rightarrow \infty$ .

It is well known that assuming **A** to **D**, if  $\pi^*(\mathbf{u}|\mathbf{R})$  is the posterior density of  $\mathbf{u} = \sqrt{T}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$ , then

$$\int (1 + \|\mathbf{u}\|^r) |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})| d\mathbf{u} \xrightarrow{P} 0, \quad 0 \leq r \leq 2, \quad (2.2)$$

where  $\phi(B, \mathbf{x})$  is the density function of a multivariate normal distribution with mean  $\mathbf{0}$  and covariance matrix  $B$  ([6, 5]).

### 2.2 Proof of Proposition 1

*Proof* Given the loss function  $\mathcal{L}(\boldsymbol{\theta}, \hat{\omega}) = h(\boldsymbol{\theta})(\hat{\omega} - \mathbf{g}(\boldsymbol{\theta}))'(\hat{\omega} - \mathbf{g}(\boldsymbol{\theta}))$ , the posterior expected value of the loss function is

$$\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{\mathcal{L}(\boldsymbol{\theta}, \hat{\omega})\} = \hat{\omega}' \hat{\omega} \mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{h(\boldsymbol{\theta})\} - 2\hat{\omega}' \mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{h(\boldsymbol{\theta})\mathbf{g}(\boldsymbol{\theta})\} + \mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{h(\boldsymbol{\theta})\mathbf{g}(\boldsymbol{\theta})'\mathbf{g}(\boldsymbol{\theta})\}$$

then

$$\frac{\partial \mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{\mathcal{L}(\boldsymbol{\theta}, \hat{\omega})\}}{\partial \hat{\omega}} = -2\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{h(\boldsymbol{\theta})\mathbf{g}(\boldsymbol{\theta})\} + 2\hat{\omega}^* \mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{h(\boldsymbol{\theta})\} = \mathbf{0}$$

so,

$$\hat{\omega}^* = \frac{\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} [\mathbf{g}(\boldsymbol{\theta})h(\boldsymbol{\theta})]}{\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} [h(\boldsymbol{\theta})]}.$$

Observe that

$$\frac{\partial^2 \mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{\mathcal{L}(\boldsymbol{\theta}, \hat{\omega})\}}{\partial \hat{\omega} \partial \hat{\omega}'} = 2\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})} \{h(\boldsymbol{\theta})\} > 0$$

### 2.3 Proof of Proposition 2

To prove 1 in Proposition 2 we have that assumptions **E** and **F** imply  $\mathbf{g}(\boldsymbol{\theta})$  and  $h(\boldsymbol{\theta})$  are constant order functions (for instance  $\mathbf{g}(\boldsymbol{\theta}) = O(1)$  and  $1/\mathbf{g}(\boldsymbol{\theta}) = O(1)$ ), then  $\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})}(\mathbf{g}(\boldsymbol{\theta})h(\boldsymbol{\theta})) = \mathbf{g}(\hat{\boldsymbol{\theta}})h(\hat{\boldsymbol{\theta}})(1 + O(T^{-1}))$  and  $\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})}(h(\boldsymbol{\theta})) = h(\hat{\boldsymbol{\theta}})(1 + O(T^{-1}))$  [7], where  $\hat{\boldsymbol{\theta}}$  is the maximum likelihood estimator. Then,

$$\begin{aligned}\hat{\omega}^* &= \frac{\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})}[\mathbf{g}(\boldsymbol{\theta})h(\boldsymbol{\theta})]}{\mathbb{E}_{\pi(\boldsymbol{\theta}|\mathbf{R})}[h(\boldsymbol{\theta})]} \\ &= \frac{\mathbf{g}(\hat{\boldsymbol{\theta}})h(\hat{\boldsymbol{\theta}})(1 + O(T^{-1}))}{h(\hat{\boldsymbol{\theta}})(1 + O(T^{-1}))} \\ &= \mathbf{g}(\hat{\boldsymbol{\theta}})(1 + O(T^{-1})) \\ &= \mathbf{g}(\hat{\boldsymbol{\theta}}) + o(1).\end{aligned}$$

To prove 2 in Proposition 2<sup>2</sup>

$$\sqrt{T}(\hat{\omega}^* - g(\boldsymbol{\theta}_0)) = \sqrt{T}(\hat{\omega}^* - g(\hat{\boldsymbol{\theta}})) + \sqrt{T}(g(\hat{\boldsymbol{\theta}}) - g(\boldsymbol{\theta}_0)),$$

Given that  $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, [I(\boldsymbol{\theta}_0)]^{-1})$ , then  $\sqrt{T}(g(\hat{\boldsymbol{\theta}}) - g(\boldsymbol{\theta}_0)) \xrightarrow{d} N(\mathbf{0}, \nabla g(\boldsymbol{\theta}_0)'[I(\boldsymbol{\theta}_0)]^{-1}g(\boldsymbol{\theta}_0))$  by the delta method. Consequently, it only remains to show that  $\sqrt{T}(\hat{\omega}^* - g(\hat{\boldsymbol{\theta}})) \xrightarrow{p} 0$  by Slutsky's theorem.

Setting  $\hat{h} = \int h(\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathbf{R})d\boldsymbol{\theta}$ ,  $\mathbf{u} = \sqrt{T}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$ , then  $\pi^*(\mathbf{u}|\mathbf{R}) = \pi\left(\frac{\mathbf{u}}{\sqrt{T}} + \hat{\boldsymbol{\theta}}|\mathbf{R}\right)\frac{1}{\sqrt{T}}$ . Using a change of variable and taking into account **E** and **F** such that we make a Taylor expansion for  $h(\boldsymbol{\theta})$

---

<sup>2</sup> We prove this for an specific weight associated with asset  $i$  to facilitate exposition. We set  $g_i(\boldsymbol{\theta}) = g(\boldsymbol{\theta})$  to simplify notation in this proof. In addition,  $c_1 = \sup_{\boldsymbol{\theta} \in \Theta} |W_N(\boldsymbol{\theta})|$  and  $c_2 = \sup_{\boldsymbol{\theta} \in \Theta} |V_N(\boldsymbol{\theta})|$ .

at  $\boldsymbol{\theta}_0$ , and for  $g(\boldsymbol{\theta})$  at  $\hat{\boldsymbol{\theta}}$ ,

$$\begin{aligned}
\sqrt{T}|\hat{\omega}^* - g(\hat{\boldsymbol{\theta}})| &= \frac{1}{\hat{h}} \left| h(\boldsymbol{\theta}_0) \int \mathbf{u}' \left[ \nabla g(\hat{\boldsymbol{\theta}}) + W_T \left( \frac{\mathbf{u}}{\sqrt{T}} + \hat{\boldsymbol{\theta}} \right) \right] \pi^*(\mathbf{u}|\mathbf{R}) d\mathbf{u} + \right. \\
&\quad \left. \frac{1}{\sqrt{T}} \int (\mathbf{u} + \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)' \left[ \nabla h(\boldsymbol{\theta}_0) + V_T \left( \frac{\mathbf{u}}{\sqrt{T}} + \hat{\boldsymbol{\theta}} \right) \right] \mathbf{u}' \times \right. \\
&\quad \left. \left[ \nabla g(\hat{\boldsymbol{\theta}}) + W_T \left( \frac{\mathbf{u}}{\sqrt{T}} + \hat{\boldsymbol{\theta}} \right) \right] \pi^*(\mathbf{u}|\mathbf{R}) d\mathbf{u} \right| \\
&\leq \frac{1}{\hat{h}} \left\{ h(\boldsymbol{\theta}_0) \left| \left\{ \int \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} \right\} \nabla g(\hat{\boldsymbol{\theta}}) \right| \right. \\
&\quad \left. + h(\boldsymbol{\theta}_0) \left| \left\{ \int \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} \right\} \mathbf{c}_1 \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| \nabla h(\boldsymbol{\theta}_0)' \left( \int \mathbf{u} \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} + \mathbf{I}(\boldsymbol{\theta}_0)^{-1} \right) \nabla g(\hat{\boldsymbol{\theta}}) \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| \mathbf{c}_2' \left( \int \mathbf{u} \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} + \mathbf{I}(\boldsymbol{\theta}_0)^{-1} \right) \nabla g(\hat{\boldsymbol{\theta}}) \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| \nabla h(\boldsymbol{\theta}_0)' \left( \int \mathbf{u} \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} + \mathbf{I}(\boldsymbol{\theta}_0)^{-1} \right) \mathbf{c}_1 \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| \mathbf{c}_2' \left( \int \mathbf{u} \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} + \mathbf{I}(\boldsymbol{\theta}_0)^{-1} \right) \mathbf{c}_1 \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)' \nabla h(\boldsymbol{\theta}_0) \int \{ \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} \} \nabla g(\hat{\boldsymbol{\theta}}) \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)' \mathbf{c}_2 \int \{ \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} \} \nabla g(\hat{\boldsymbol{\theta}}) \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)' \nabla h(\boldsymbol{\theta}_0) \int \{ \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} \} \mathbf{c}_1 \right| \right. \\
&\quad \left. + \frac{1}{\sqrt{T}} \left| (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)' \mathbf{c}_2 \int \{ \mathbf{u}' (\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\boldsymbol{\theta}_0)^{-1}, \mathbf{u})) d\mathbf{u} \} \mathbf{c}_1 \right| \right\},
\end{aligned}$$

then,

$$\begin{aligned}
\sqrt{T}(\hat{\omega}^* - g(\hat{\theta})) &\leq \frac{1}{\hat{h}} \left\{ h(\theta_0) \|\nabla g(\hat{\theta})\| \left\{ \int \|\mathbf{u}\| |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} \right\} \right. \\
&\quad + h(\theta_0) c_1 \left\{ \int \|\mathbf{u}\| |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} \right\} \\
&\quad + \frac{1}{\sqrt{T}} \|\nabla h(\theta_0)\| \|\nabla g(\hat{\theta})\| \left( \int \|\mathbf{u}\|^2 |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} + \mathbf{I}(\theta_0)^{-1} \right) \\
&\quad + \frac{1}{\sqrt{T}} c_2 \|\nabla g(\hat{\theta})\| \left( \int \|\mathbf{u}\|^2 |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} + \mathbf{I}(\theta_0)^{-1} \right) \\
&\quad + \frac{1}{\sqrt{T}} \|\nabla h(\theta_0)\| c_1 \left( \int \|\mathbf{u}\|^2 |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} + \mathbf{I}(\theta_0)^{-1} \right) \\
&\quad + \frac{1}{\sqrt{T}} c_2 c_1 \left( \int \|\mathbf{u}\|^2 |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} + \mathbf{I}(\theta_0)^{-1} \right) \\
&\quad + \frac{1}{\sqrt{T}} \|\hat{\theta} - \theta_0\| \|\nabla h(\theta_0)\| \|\nabla g(\hat{\theta})\| \int \|\mathbf{u}\| |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} \\
&\quad + \frac{1}{\sqrt{T}} \|\hat{\theta} - \theta_0\| c_2 \|\nabla g(\hat{\theta})\| \int \|\mathbf{u}\| |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} \\
&\quad + \frac{1}{\sqrt{T}} \|\hat{\theta} - \theta_0\| \|\nabla h(\theta_0)\| c_1 \int \|\mathbf{u}\| |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} \\
&\quad \left. + \frac{1}{\sqrt{T}} \|\hat{\theta} - \theta_0\| c_2 c_1 \int \|\mathbf{u}\| |\pi^*(\mathbf{u}|\mathbf{R}) - \phi(\mathbf{I}(\theta_0)^{-1}, \mathbf{u})| d\mathbf{u} \right\},
\end{aligned}$$

taking into account that  $\hat{h} \xrightarrow{P} h(\theta_0) \neq 0$ ,  $\hat{\theta} \xrightarrow{P} \theta_0$ , assumptions **A**(4), **E**, **F** and result 2.2, then we conclude that  $\sqrt{T}(\hat{\omega}^* - g(\hat{\theta})) \xrightarrow{P} 0$  by Theorem 2.1.3 in [5].

□

## 2.4 Simulation exercises

### 2.4.1 Simulation details

Tables 2.1 and 2.2 show the mean vectors and covariance matrices of excess of returns for global minimum variance and tangency portfolios simulation exercises using 10 and 50 stocks, respectively.

**Table 2.1:** Mean and correlation matrix: 10 assets

Expected Return	Correlation matrix									
-12,39%	1,0000	0,0931	0,1120	0,1306	-0,0967	0,0982	-0,0402	-0,1942	0,0469	-0,1207
16,43%	-	1,0000	0,0764	0,0523	0,0309	-0,0450	-0,0605	-0,0065	0,1226	0,1290
-10,89%	-	-	1,0000	0,0022	0,0718	0,1121	-0,0317	0,0096	0,0877	-0,1039
13,00%	-	-	-	1,0000	0,0817	0,0479	0,1425	-0,0614	0,1845	-0,0765
16,62%	-	-	-	-	1,0000	0,0641	-0,1029	0,2510	-0,0741	0,0039
0,21%	-	-	-	-	-	1,0000	-0,0998	-0,0420	-0,0568	-0,0917
3,76%	-	-	-	-	-	-	1,0000	0,0461	-0,0186	0,0285
-9,34%	-	-	-	-	-	-	-	1,0000	-0,0519	-0,0361
1,73%	-	-	-	-	-	-	-	-	1,0000	-0,0069
18,75%	-	-	-	-	-	-	-	-	-	1,0000



parameters imply a non-informative prior distribution for  $\Sigma$ . This implies uncertainty regarding  $\Omega$ .<sup>3</sup>

**Table 2.3:**  $\alpha$ ,  $\beta$ , and  $H$ : 10 assets

Alpha	Betas	Error variance									
-2,15%	0,9028	3,8837	1,7186	-2,2844	-0,4246	-3,9635	-2,1763	-1,6144	-1,4926	-0,9370	-1,0180
8,66%	-0,7351	-	4,1537	-0,7808	-0,8591	-1,6083	0,2882	0,3573	-1,2962	-0,5773	-1,1712
-1,58%	0,2621	-	-	5,4638	1,2925	3,7016	-0,4715	-2,0064	2,7301	-1,6119	1,3410
7,37%	-1,2975	-	-	-	5,8486	0,8501	0,1784	0,3940	1,9296	-1,8786	1,3972
8,73%	-0,8429	-	-	-	-	5,5818	1,7485	1,4281	2,7956	0,7115	0,4547
2,58%	0,5402	-	-	-	-	-	7,1260	4,5114	2,7744	0,7845	0,6050
3,91%	-1,3512	-	-	-	-	-	-	6,1486	0,1098	1,6010	0,2154
-1,00%	0,1004	-	-	-	-	-	-	-	5,8499	-0,5968	1,3032
3,15%	-0,7126	-	-	-	-	-	-	-	-	2,1748	-0,8852
9,53%	1,6904	-	-	-	-	-	-	-	-	-	2,6829

<sup>3</sup> We only use informative priors for location parameters in tangency portfolio exercises. In addition, we only use informative priors for  $\alpha$  in Treynor–Black exercises.

Table 2.4:  $\alpha$ ,  $\beta$ , and  $H$ : 50 assets

Asset	$\alpha$	$\beta$	$H$
1	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0000
7	0.0000	0.0000	0.0000
8	0.0000	0.0000	0.0000
9	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000
11	0.0000	0.0000	0.0000
12	0.0000	0.0000	0.0000
13	0.0000	0.0000	0.0000
14	0.0000	0.0000	0.0000
15	0.0000	0.0000	0.0000
16	0.0000	0.0000	0.0000
17	0.0000	0.0000	0.0000
18	0.0000	0.0000	0.0000
19	0.0000	0.0000	0.0000
20	0.0000	0.0000	0.0000
21	0.0000	0.0000	0.0000
22	0.0000	0.0000	0.0000
23	0.0000	0.0000	0.0000
24	0.0000	0.0000	0.0000
25	0.0000	0.0000	0.0000
26	0.0000	0.0000	0.0000
27	0.0000	0.0000	0.0000
28	0.0000	0.0000	0.0000
29	0.0000	0.0000	0.0000
30	0.0000	0.0000	0.0000
31	0.0000	0.0000	0.0000
32	0.0000	0.0000	0.0000
33	0.0000	0.0000	0.0000
34	0.0000	0.0000	0.0000
35	0.0000	0.0000	0.0000
36	0.0000	0.0000	0.0000
37	0.0000	0.0000	0.0000
38	0.0000	0.0000	0.0000
39	0.0000	0.0000	0.0000
40	0.0000	0.0000	0.0000
41	0.0000	0.0000	0.0000
42	0.0000	0.0000	0.0000
43	0.0000	0.0000	0.0000
44	0.0000	0.0000	0.0000
45	0.0000	0.0000	0.0000
46	0.0000	0.0000	0.0000
47	0.0000	0.0000	0.0000
48	0.0000	0.0000	0.0000
49	0.0000	0.0000	0.0000
50	0.0000	0.0000	0.0000

Tables 2.5 and 2.6 show optimal population weights given previous setting.



**Table 2.5:** Optimal population weights for tangency, global minimum variance, and Treynor-Black portfolios: 10 assets

Weights		
TP	GMV	TB
-48,02%	12,11%	32,58%
54,99%	6,72%	5,56%
-36,51%	7,63%	-4,15%
41,20%	5,63%	8,12%
65,24%	8,74%	29,49%
9,84%	11,57%	14,25%
16,21%	12,24%	-17,20%
-53,54%	12,00%	-14,45%
0,41%	9,76%	23,31%
50,17%	13,61%	22,49%

TP: Tangency portfolio, GMV: Global minimum variance, and TB: Treynor-Black

**Table 2.6:** Optimal population weights for tangency, global minimum variance, and Treynor-Black portfolios: 50 assets

TP	MGV	TB
-15,91%	2,07%	31,77%
34,93%	3,91%	-0,58%
2,70%	2,02%	9,24%
10,05%	1,95%	6,57%
13,45%	1,52%	-0,05%
7,72%	4,46%	7,36%
-2,29%	1,13%	8,97%
-1,76%	5,84%	3,35%
1,31%	0,77%	-0,09%
21,79%	3,43%	2,80%
22,85%	0,59%	8,96%
1,71%	-1,06%	1,70%
-4,79%	0,22%	25,16%
-31,67%	2,88%	-11,61%
4,76%	5,12%	25,81%
8,98%	3,95%	-10,13%
1,19%	1,48%	33,44%
-5,71%	2,73%	-1,01%
27,90%	2,50%	-23,76%
1,16%	2,15%	2,88%
21,58%	1,28%	-2,18%
-28,95%	1,18%	9,18%
-9,84%	-1,87%	-6,27%
-27,98%	3,22%	13,68%
16,28%	2,13%	8,10%
-1,34%	3,65%	18,29%
-2,50%	1,20%	7,44%
0,59%	1,77%	-28,79%
-8,26%	0,66%	-12,95%
-1,32%	4,09%	-3,05%
-6,38%	2,72%	0,22%
-20,59%	-0,02%	-10,91%
4,50%	3,38%	-7,49%
-6,28%	4,91%	-21,41%
-27,90%	-0,65%	0,76%
13,55%	3,11%	5,30%
7,78%	2,09%	-4,07%
17,94%	1,04%	16,65%
42,64%	3,34%	13,19%
-31,53%	1,09%	-20,80%
7,17%	2,61%	-8,57%
-4,28%	2,64%	8,75%
26,14%	1,67%	-9,71%
-21,12%	0,64%	-17,58%
-15,29%	2,03%	4,91%
37,61%	1,07%	43,31%
3,75%	0,97%	-12,75%
-0,90%	0,32%	1,49%
-7,94%	0,35%	-3,92%
24,51%	1,73%	-1,58%

TP: Tangency portfolio, GMV: Global minimum variance, and TB: Treynor-Black

Table 2.7 the optimal values of the different objective functions.

**Table 2.7:** Objective Values

portfolio Size	Sharpe Ratio	Variance	Inf. Ratio
10	0,3793	0,1060	0,3214
50	1,2081	0.0150	1,9090

#### 2.4.2 Robustness Check

The previous simulation exercises assume that returns are serially independent and normal distributed. However, these assumption may be not satisfied in practice. Therefore, we conduct some simulation exercises to check robustness of our results.

We set population parameters as in the main framework but stochastic errors are drawn from a multivariate *Student's t*-distribution with 4 degrees of freedom or *AR*(1) process.

Table 2.8 and Table 2.9 show the summary statistics for the squared errors of global minimum variance portfolio weights when the data is generated from a multivariate *Student's t*-distribution and *AR*(1) process, respectively. We can observe that there are not significant differences. We got similar results in the main setting. As a consequence of these results, we do not observe any substantial difference in variance estimates (Tables 2.10 and 2.11).

Table 2.12 and Table 2.13 show summary statistics of squared errors for tangency portfolio using a multivariate *Student's-t* or *AR*(1) process to generate the data. We observe that our MELO proposal have the lowest degree of variability as well as MSE. Tables 2.14 and 2.15 show the descriptive statistics of Sharpe ratios. Informative approaches dominate estimation performance. Finally, the Treynor–Black model present similar results as the initial simulation. Informative MELO shows the lowest MSE and plug-in the highest MSE in all cases (Tables 2.16 and 2.17). Regarding information ratio, there is no significant difference between any of the methodologies, although the informative cases present a little improvement (Tables 2.18 and 2.19). In summary, the good results of our MELO proposal that are presented in the main framework are robust to heavy tails and serial dependence.

**Table 2.8:** Weights: Global minimum variance using Student's-t

Stocks = 10							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.0014	0.0110	0.0164	0.0191	0.0243	0.0727
Bayesian NI		0.0014	0.0110	0.0164	0.0191	0.0243	0.0727
Bayesian I	120	0.0015	0.0106	0.0159	0.0187	0.0239	0.0716
MELO NI		0.0016	0.0111	0.0164	0.0192	0.0243	0.0731
MELO I		0.0011	0.0105	0.0161	0.0184	0.0236	0.0736
Naive ( $1/N$ )		0.0066	0.0066	0.0066	0.0066	0.0066	0.0066
Plug-in		0.0006	0.0060	0.0088	0.0101	0.0125	0.0668
Bayesian NI		0.0006	0.0060	0.0088	0.0101	0.0125	0.0668
Bayesian I	240	0.0007	0.0060	0.0087	0.0100	0.0124	0.0665
MELO NI		0.0006	0.0061	0.0087	0.0102	0.0125	0.0656
MELO I		0.0011	0.0063	0.0089	0.0103	0.0128	0.0620
Naive ( $1/N$ )		0.0023	0.0023	0.0023	0.0023	0.0023	0.0023
Stocks = 50							
Plug-in		0.0084	0.0182	0.0223	0.0234	0.0276	0.0553
Bayesian NI		0.0084	0.0182	0.0223	0.0234	0.0276	0.0553
Bayesian I	120	0.0080	0.0170	0.0208	0.0219	0.0259	0.0526
MELO NI		0.0089	0.0184	0.0224	0.0235	0.0278	0.0560
MELO I		0.0075	0.0178	0.0215	0.0225	0.0263	0.0528
Naive ( $1/N$ )		0.0123	0.0123	0.0123	0.0123	0.0123	0.0123
Plug-in		0.0033	0.0068	0.0081	0.0083	0.0095	0.0219
Bayesian NI		0.0033	0.0068	0.0081	0.0083	0.0095	0.0219
Bayesian I	240	0.0032	0.0067	0.0079	0.0082	0.0094	0.0214
MELO NI		0.0035	0.0069	0.0081	0.0084	0.0096	0.0224
MELO I		0.0037	0.0071	0.0084	0.0087	0.0099	0.0208
Naive ( $1/N$ )		0.0056	0.0056	0.0056	0.0056	0.0056	0.0056

Note: This table shows summary statistics for the squared error.  
 "NI": Non-informative priors. "I": Informative priors. Stochastic errors  
 are drawn from a multivariate student's-t distribution with 4 degrees of freedom.

**Table 2.9:** Weights: Global minimum variance using AR(1)

Stocks = 10							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.0010	0.0072	0.0102	0.0109	0.0137	0.0337
Bayesian NI		0.0010	0.0072	0.0102	0.0109	0.0137	0.0337
Bayesian I	120	0.0013	0.0072	0.0101	0.0108	0.0135	0.0329
MELO NI		0.0008	0.0074	0.0103	0.0110	0.0138	0.0325
MELO I		0.0009	0.0076	0.0106	0.0112	0.0139	0.0308
Naive (1/N)		0.0066	0.0066	0.0066	0.0066	0.0066	0.0066
Plug-in		0.0017	0.0069	0.0091	0.0095	0.0118	0.0249
Bayesian NI		0.0017	0.0069	0.0091	0.0095	0.0118	0.0249
Bayesian I	240	0.0017	0.0070	0.0092	0.0096	0.0118	0.0254
MELO NI		0.0017	0.0069	0.0092	0.0095	0.0119	0.0243
MELO I		0.0024	0.0076	0.0099	0.0103	0.0125	0.0260
Naive (1/N)		0.0023	0.0023	0.0023	0.0023	0.0023	0.0023
Stocks = 50							
Plug-in		0.0066	0.0160	0.0195	0.0202	0.0238	0.0419
Bayesian NI		0.0066	0.0160	0.0195	0.0202	0.0238	0.0419
Bayesian I	120	0.0064	0.0154	0.0187	0.0194	0.0227	0.0416
MELO NI		0.0066	0.0161	0.0196	0.0203	0.0238	0.0414
MELO I		0.0068	0.0158	0.0191	0.0197	0.0230	0.0436
Naive (1/N)		0.0123	0.0123	0.0123	0.0123	0.0123	0.0123
Plug-in		0.0038	0.0066	0.0078	0.0079	0.0089	0.0154
Bayesian NI		0.0038	0.0066	0.0078	0.0079	0.0089	0.0154
Bayesian I	240	0.0038	0.0065	0.0077	0.0078	0.0088	0.0151
MELO NI		0.0038	0.0066	0.0078	0.0079	0.009	0.0152
MELO I		0.0039	0.007	0.0082	0.0083	0.0094	0.0155
Naive (1/N)		0.0056	0.0056	0.0056	0.0056	0.0056	0.0056

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors.

Stochastic errors are drawn from a AR(1) process.

**Table 2.10:** Variance: Global minimum variance using Student's-t

Stocks = 10; sigma_120= 0.325723; sigma_120=0.3212232							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.3280	0.3410	0.3484	0.3513	0.3579	0.4125
Bayesian NI		0.3280	0.3410	0.3484	0.3513	0.3579	0.4125
Bayesian I	120	0.3284	0.3406	0.3478	0.3507	0.3574	0.4091
MELO NI		0.3280	0.3410	0.3484	0.3514	0.3581	0.4166
MELO I		0.3276	0.3404	0.3475	0.3505	0.3571	0.4131
Naive (1/N)		0.3349	0.3349	0.3349	0.3349	0.3349	0.3349
Plug-in		0.3223	0.3301	0.3341	0.3360	0.3395	0.4060
Bayesian NI		0.3223	0.3301	0.3341	0.3360	0.3395	0.4060
Bayesian I	240	0.3224	0.3300	0.3340	0.3359	0.3393	0.4053
MELO NI		0.3223	0.3302	0.3343	0.3361	0.3396	0.4046
MELO I		0.3230	0.3305	0.3344	0.3363	0.3400	0.4005
Naive (1/N)		0.3247	0.3247	0.3247	0.3247	0.3247	0.3247
Stocks = 50; sigma_120=0.122543 ; sigma_240=0.1253086							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.1450	0.1641	0.1713	0.1724	0.1796	0.2145
Bayesian NI		0.1450	0.1641	0.1713	0.1724	0.1796	0.2145
Bayesian I	120	0.1432	0.1619	0.1685	0.1695	0.1765	0.2086
MELO NI		0.1455	0.1643	0.1715	0.1726	0.1798	0.2141
MELO I		0.1433	0.1628	0.1694	0.1704	0.1770	0.2130
Naive (1/N)		0.1509	0.1509	0.1509	0.1509	0.1509	0.1509
Plug-in		0.1379	0.1462	0.1494	0.1500	0.1535	0.1756
Bayesian NI		0.1379	0.1462	0.1494	0.1500	0.1535	0.1756
Bayesian I	240	0.1376	0.1457	0.1490	0.1496	0.1530	0.1744
MELO NI		0.1378	0.1464	0.1494	0.1501	0.1536	0.1761
MELO I		0.1379	0.1468	0.1503	0.1508	0.1543	0.1755
Naive (1/N)		0.1422	0.1422	0.1422	0.1422	0.1422	0.1422

Note: This table shows summary statistics for the squared error.  
 "NI": Non-informative priors. "I": Informative priors. Stochastic errors  
 are drawn from a multivariate student's-t distribution with 4 degrees of freedom.

**Table 2.11:** Variance: Global minimum variance using AR(1)

Stocks = 10; sigma_120= 0.325723; sigma_120=0.3212232							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.3272	0.3360	0.3399	0.3409	0.3446	0.3748
Bayesian NI		0.3272	0.3360	0.3399	0.3409	0.3446	0.3748
Bayesian I	120	0.3275	0.3359	0.3400	0.3407	0.3445	0.3737
MELO NI		0.3272	0.3361	0.3401	0.3410	0.3447	0.3732
MELO I		0.3271	0.3362	0.3404	0.3412	0.3452	0.3688
Naive (1/N)		0.3349	0.3349	0.3349	0.3349	0.3349	0.3349
Plug-in		0.3240	0.3312	0.3343	0.3348	0.3381	0.3545
Bayesian NI		0.3240	0.3312	0.3343	0.3348	0.3381	0.3545
Bayesian I	240	0.3240	0.3313	0.3344	0.3349	0.3381	0.3546
MELO NI		0.3240	0.3312	0.3343	0.3349	0.3381	0.3537
MELO I		0.3249	0.3322	0.3355	0.3359	0.3391	0.3560
Naive (1/N)		0.3247	0.3247	0.3247	0.3247	0.3247	0.3247
Stocks = 50; sigma_120=0.122543 ; sigma_240=0.1253086							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.1420	0.1606	0.1664	0.1672	0.1730	0.2037
Bayesian NI		0.1420	0.1606	0.1664	0.1672	0.1730	0.2037
Bayesian I	120	0.1400	0.1591	0.1648	0.1655	0.1711	0.2038
MELO NI		0.1421	0.1608	0.1666	0.1674	0.1733	0.2036
MELO I		0.1441	0.1597	0.1658	0.1665	0.1719	0.2102
Naive (1/N)		0.1509	0.1509	0.1509	0.1509	0.1509	0.1509
Plug-in		0.1381	0.1459	0.1491	0.1493	0.1522	0.1661
Bayesian NI		0.1381	0.1459	0.1491	0.1493	0.1522	0.1661
Bayesian I	240	0.1374	0.1458	0.1489	0.1490	0.1519	0.1657
MELO NI		0.1380	0.1462	0.1491	0.1494	0.1523	0.1661
MELO I		0.1374	0.1467	0.1495	0.1499	0.1528	0.1667
Naive (1/N)		0.1422	0.1422	0.1422	0.1422	0.1422	0.1422

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors.

Stochastic errors are drawn from a AR(1) process.

**Table 2.12:** Weights: Tangency portfolio using Student's-t

Stocks = 10								
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max	
Plug-in		0.0520	0.7813	1.2893	5,259.6005	7.6138	3,602,104.8449	
Shrinkage		0.0856	0.8145	1.2034	2,526.7271	4.2920	1,618,338.6448	
Bayesian NI	120	0.0520	0.7813	1.2893	5,259.6005	7.6138	3,602,104.8449	
Bayesian I		0.0294	0.3862	0.5967	6,605.7032	1.8151	5,975,170.4390	
MELO NI		0.1323	0.7493	0.9746	1.3971	1.3981	7.5181	
MELO I		0.0409	0.3281	0.4665	0.8212	0.6648	11.9204	
Naive (1/N)		1.7360	1.7360	1.7360	1.7360	1.7360	1.7360	
Plug-in			0.0457	0.3059	0.4915	124.2089	1.4563	50,923.1849
Shrinkage			0.0657	0.3178	0.4681	82.5681	0.8574	35,360.7469
Bayesian NI	240	0.0457	0.3059	0.4915	124.2089	1.4563	50,923.1849	
Bayesian I		0.0155	0.1379	0.2122	10.6512	0.4132	7,682.8385	
MELO NI		0.0646	0.2640	0.3560	0.5252	0.5027	6.3474	
MELO I		0.0212	0.1322	0.2008	0.3483	0.3201	14.0347	
Naive (1/N)		1.1570	1.1570	1.1570	1.1570	1.1570	1.1570	
Plug-in			0.2911	0.7799	1.2011	14,808.0309	7.9732	14,571,788.8158
Shrinkage			0.2924	0.7625	0.9772	11,432.4759	4.4684	11,297,502.1113
Bayesian NI	120	0.2911	0.7799	1.2011	14,808.0309	7.9732	14,571,788.8155	
Bayesian I		0.2544	0.6139	0.8567	6,145.2737	4.5235	4,961,375.0658	
MELO NI		0.2818	0.6739	0.8205	1.3409	1.0479	7.9691	
MELO I		0.4558	0.7586	0.8497	1.0169	0.9972	3.6699	
Naive (1/N)		1.5310	1.5310	1.5310	1.5310	1.5310	1.5310	
Plug-in			0.1536	0.2772	0.3685	1,302.9453	1.1773	824,784.4664
Shrinkage			0.1429	0.2667	0.3261	752.5638	0.7324	438,595.0858
Bayesian NI	240	0.1536	0.2772	0.3685	1,302.9453	1.1773	824,784.4664	
Bayesian I		0.1056	0.2048	0.2695	323.4630	0.6132	286,826.2194	
MELO NI		0.1317	0.2519	0.2971	0.4361	0.3650	4.7133	
MELO I		0.1076	0.1877	0.2179	0.2649	0.2547	2.9808	
Naive (1/N)		0.7058	0.7058	0.7058	0.7058	0.7058	0.7058	

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors. Stochastic errors are drawn from a multivariate student's-t distribution with 4 degrees of freedom.

**Table 2.13:** Weights: Tangency portfolio using AR(1)

Stocks = 10							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.2443	1.0457	2.2594	2,111.7934	11.3143	1,463,010.2553
Shrinkage		0.2689	1.0409	1.7287	1,066.0852	8.5640	645,298.1837
Bayesian NI		0.2443	1.0457	2.2594	2,111.7934	11.3143	1,463,010.2553
Bayesian I	120	0.0644	0.4017	0.6626	1,640.8696	3.0662	455,446.3691
MELO NI		0.2085	0.9281	1.2291	1.6885	2.0192	6.4059
MELO I		0.0365	0.3386	0.4869	0.9640	0.6665	11.0816
Naive (1/N)		1.7360	1.7360	1.7360	1.7360	1.7360	1.7360
Plug-in		0.0494	0.4354	1.0767	1,845,443.8977	9.6919	1,843,392,342.3385
Shrinkage		0.0638	0.4165	0.8586	1,760,248.3245	7.2455	1,758,350,678.9414
Bayesian NI		0.0494	0.4354	1.0767	1,845,443.8978	9.6919	1,843,392,342.4105
Bayesian I	240	0.0311	0.1646	0.3364	3,562.5269	1.3584	2,940,030.6373
MELO NI		0.0676	0.3278	0.4984	0.9248	0.8190	6.2356
MELO I		0.0229	0.1442	0.2549	0.5990	0.6372	15.0396
Naive (1/N)		1.1570	1.1570	1.1570	1.1570	1.1570	1.1570
Stocks = 50							
Plug-in		0.5020	0.9803	1.3303	747.0103	5.6053	216,372.6336
Shrinkage		0.4953	0.9592	1.2008	513.7773	4.3128	161,295.0070
Bayesian NI		0.5020	0.9803	1.3303	747.0103	5.6053	216,372.6336
Bayesian I	120	0.3408	0.6307	0.7530	129.3772	1.0373	55,913.0072
MELO NI		0.4081	0.8791	1.0139	1.2973	1.2336	6.5596
MELO I		0.4523	0.7012	0.7841	0.8544	0.8835	3.3046
Naive (1/N)		1.5310	1.5310	1.5310	1.5310	1.5310	1.5310
Plug-in		0.165	0.3372	0.4488	69.5874	1.0824	27,495.787
Shrinkage		0.1603	0.3319	0.4127	53.6413	0.8294	20,424.9287
Bayesian NI		0.165	0.3372	0.4488	69.5874	1.0824	27,495.787
Bayesian I	240	0.091	0.1796	0.2143	67.8975	0.2864	56,790.8303
MELO NI		0.1559	0.3084	0.3645	0.4529	0.4384	3.6619
MELO I		0.0995	0.1764	0.2037	0.2126	0.2329	1.5794
Naive (1/N)		0.7058	0.7058	0.7058	0.7058	0.7058	0.7058

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors.

Stochastic errors are drawn from a AR(1) process.



**Table 2.14:** Sharpe ratio: Tangency portfolio using Student's-t

Stocks= 10; SR_120=0.379298; SR_240=0.3866669							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.3475	0.2364	0.2877	0.1948	0.3152	0.3728
Shrinkage		-0.3469	0.2145	0.2802	0.1883	0.3120	0.3702
Bayesian NI		-0.3475	0.2364	0.2877	0.1948	0.3152	0.3728
Bayesian I	120	-0.3589	0.3281	0.3433	0.3055	0.3540	0.3766
MELO NI		-0.2938	0.2126	0.2804	0.2031	0.3106	0.3706
MELO I		-0.3395	0.3263	0.3431	0.3081	0.3544	0.3772
Naive (1/N)		0.1131	0.1131	0.1131	0.1131	0.1131	0.1131
Plug-in		-0.3508	0.3143	0.3355	0.3006	0.3496	0.3799
Shrinkage		-0.3491	0.3116	0.3340	0.2988	0.3482	0.3802
Bayesian NI		-0.3508	0.3143	0.3355	0.3006	0.3496	0.3799
Bayesian I	240	-0.3436	0.3573	0.3654	0.3617	0.3717	0.3844
MELO NI		-0.3212	0.3116	0.3350	0.3019	0.3497	0.3787
MELO I		-0.3336	0.3578	0.3655	0.3620	0.3718	0.3839
Naive (1/N)		0.1167	0.1167	0.1167	0.1167	0.1167	0.1167
Stocks= 50; SR_120=1.20819; SR_240=0.9470014							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.9154	0.6681	0.7485	0.5177	0.8073	1.0179
Shrinkage		-0.9060	0.6599	0.7409	0.5116	0.8024	1.0129
Bayesian NI		-0.9154	0.6681	0.7485	0.5177	0.8073	1.0179
Bayesian I	120	-0.9783	0.7923	0.8431	0.6777	0.8868	1.0324
MELO NI		-0.8946	0.6496	0.7384	0.5210	0.7984	1.0226
MELO I		-0.8273	0.6926	0.7820	0.6271	0.8318	0.9754
Naive (1/N)		0.1188	0.1188	0.1188	0.1188	0.1188	0.1188
Plug-in		-0.8150	0.7013	0.7294	0.6560	0.7592	0.8420
Shrinkage		-0.8146	0.7000	0.7279	0.6550	0.7590	0.8440
Bayesian NI		-0.8150	0.7013	0.7294	0.6560	0.7592	0.8420
Bayesian I	240	-0.8589	0.7648	0.7881	0.7409	0.8079	0.8788
MELO NI		-0.7553	0.6990	0.7286	0.6550	0.7583	0.8523
MELO I		-0.8036	0.7572	0.7815	0.7372	0.8035	0.8718
Naive (1/N)		0.1260	0.1260	0.1260	0.1260	0.1260	0.1260

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors. Stochastic errors are drawn from a multivariate student's-t distribution with 4 degrees of freedom.

**Table 2.15:** Sharpe ratio: Tangency portfolio using AR(1)

Stocks= 10; SR_120=0.379298; SR_240=0.3866669							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.3396	-0.0526	0.2452	0.1370	0.2954	0.3595
Shrinkage		-0.3393	-0.0517	0.2441	0.1370	0.2932	0.3567
Bayesian NI		-0.3396	-0.0526	0.2452	0.1370	0.2954	0.3595
Bayesian I	120	-0.3562	0.3253	0.3413	0.2865	0.3528	0.3750
MELO NI		-0.2910	0.0547	0.2358	0.1528	0.2866	0.3613
MELO I		-0.3330	0.3236	0.3414	0.2911	0.3531	0.3765
Naive (1/N)		0.1131	0.1131	0.1131	0.1131	0.1131	0.1131
Plug-in		-0.3538	0.2516	0.3119	0.2126	0.3380	0.3804
Shrinkage		-0.3537	0.2507	0.3116	0.2124	0.3377	0.3795
Bayesian NI		-0.3538	0.2516	0.3119	0.2126	0.3380	0.3804
Bayesian I	240	-0.3505	0.3548	0.3636	0.3520	0.3713	0.3837
MELO NI		-0.3174	0.2450	0.3095	0.2242	0.3377	0.3784
MELO I		-0.3380	0.3552	0.3643	0.3530	0.3718	0.3839
Naive (1/N)		0.1167	0.1167	0.1167	0.1167	0.1167	0.1167
Stocks= 50; SR_120=1.20819; SR_240=0.9470014							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.7976	0.5276	0.6224	0.4656	0.6975	0.9057
Shrinkage		-0.7948	0.5280	0.6223	0.4659	0.6955	0.9052
Bayesian NI		-0.7976	0.5276	0.6224	0.4656	0.6975	0.9057
Bayesian I	120	-0.9322	0.8002	0.8358	0.7750	0.8766	0.9993
MELO NI		-0.7309	0.5235	0.6219	0.4795	0.6933	0.9164
MELO I		-0.7833	0.7728	0.8158	0.7580	0.8548	0.9927
Naive (1/N)		0.1188	0.1188	0.1188	0.1188	0.1188	0.1188
Plug-in		-0.7454	0.6335	0.6719	0.6049	0.7123	0.8275
Shrinkage		-0.7450	0.6346	0.6720	0.6050	0.7121	0.8295
Bayesian NI		-0.7454	0.6335	0.6719	0.6049	0.7123	0.8275
Bayesian I	240	-0.8014	0.7870	0.8054	0.7974	0.8210	0.8685
MELO NI		-0.7076	0.6307	0.6716	0.6116	0.7116	0.8259
MELO I		-0.6929	0.7838	0.8016	0.7929	0.8181	0.8746
Naive (1/N)		0.1260	0.1260	0.1260	0.1260	0.1260	0.1260

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors.

Stochastic errors are drawn from a AR(1) process.

**Table 2.16:** Weights: Treynor-Black using Student's-t

Stocks = 10							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.2840	0.8241	1.7772	465.7448	6.2597	123,955.8744
Bayesian NI		0.0016	0.0286	0.0669	336.2529	0.2235	158,033.8534
Bayesian I	120	0.0005	0.0066	0.0169	0.0760	0.0433	10.2129
MELO NI		0.0021	0.0182	0.0420	0.0885	0.0885	1.6260
MELO I		0.0007	0.0065	0.0154	0.0272	0.0312	0.5434
Naive (1/N)		0.2802	0.2802	0.2802	0.2802	0.2802	0.2802
Plug-in		0.2372	0.6706	1.2859	1,254.1817	4.3356	1,039,877.1002
Bayesian NI		0.0011	0.0135	0.0343	146.9521	0.0959	144,125.6329
Bayesian I	240	0.0003	0.0035	0.0081	0.0235	0.0221	0.9403
MELO NI		0.0015	0.0118	0.0257	0.0455	0.0562	1.5436
MELO I		0.0003	0.0033	0.0089	0.0170	0.0198	0.4438
Naive (1/N)		0.2802	0.2802	0.2802	0.2802	0.2802	0.2802
Stocks = 50							
Plug-in		0.9544	1.2917	1.8874	5,557.4684	4.7195	4,267,788.4470
Bayesian NI		0.0102	0.1403	0.3039	0.5308	0.6233	12.6842
Bayesian I	120	0.0013	0.0061	0.0116	0.0224	0.0253	0.5972
MELO NI		0.0011	0.0072	0.0145	0.0282	0.0298	0.6366
MELO I		0.0013	0.0060	0.0115	0.0217	0.0250	0.4954
Naive (1/N)		1.0380	1.0380	1.0380	1.0380	1.0380	1.0380
Plug-in		0.9059	1.1102	1.3273	340.1823	2.3147	117,699.5531
Bayesian NI		0.0129	0.1593	0.2646	0.3363	0.4349	2.4870
Bayesian I	240	0.0005	0.0029	0.0052	0.0095	0.0119	0.0968
MELO NI		0.0008	0.0033	0.0063	0.0111	0.0133	0.0986
MELO I		0.0004	0.0029	0.0052	0.0094	0.0116	0.0920
Naive (1/N)		1.0380	1.0380	1.0380	1.0380	1.0380	1.0380

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors. Stochastic errors are drawn from a multivariate student's-t distribution with 4 degrees of freedom.

**Table 2.17:** Weights: Treynor-Black using AR(1)

Stocks = 10							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		0.2842	0.9042	1.9296	1,259.3917	7.9150	622,875.1781
Bayesian NI		0.0386	0.4883	1.6043	212.1123	7.5663	98,674.7024
Bayesian I	120	0.0313	0.3451	1.0556	1,032.9164	4.3053	699,137.6630
MELO NI		0.0151	0.1747	0.3005	0.4110	0.5126	3.3628
MELO I		0.0282	0.0957	0.1395	0.2062	0.2118	2.3360
Naive (1/N)		0.2802	0.2802	0.2802	0.2802	0.2802	0.2802
Plug-in		0.2117	0.7332	1.4472	1,007.6181	5.2083	608,862.9270
Bayesian NI		0.0330	0.3761	1.6678	2,369.7168	9.5485	1,671,944.1212
Bayesian I	240	0.0730	0.4133	0.9935	67,889.4698	2.9586	67,811,745.6635
MELO NI		0.0154	0.1338	0.2276	0.3704	0.4464	4.3305
MELO I		0.0389	0.1178	0.1678	0.2101	0.2449	2.0109
Naive (1/N)		0.2802	0.2802	0.2802	0.2802	0.2802	0.2802
Stocks = 50							
Plug-in		0.9153	1.4745	2.3718	6,021.6713	6.9250	5,525,522.2559
Bayesian NI		0.3203	0.9784	1.4347	1,581.8523	3.4823	1,322,490.9264
Bayesian I	120	0.3534	0.7059	0.8081	1.0619	0.9075	240.7099
MELO NI		0.3384	0.8726	1.0544	1.1292	1.2917	3.2231
MELO I		0.4022	0.7142	0.8041	0.8081	0.8957	1.2478
Naive (1/N)		1.0380	1.0380	1.0380	1.0380	1.0380	1.0380
Plug-in		0.9080	1.2728	1.8210	39,426.6828	4.5423	38,989,535.9672
Bayesian NI		0.3988	0.8340	1.0559	18.8188	1.4966	5,449.2732
Bayesian I	240	0.4497	0.7007	0.7765	0.7735	0.8496	1.1393
MELO NI		0.3688	0.8017	0.9573	0.9812	1.1197	2.4815
MELO I		0.4693	0.7108	0.7827	0.7789	0.8503	1.0849
Naive (1/N)		1.0380	1.0380	1.0380	1.0380	1.0380	1.0380

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors.

Stochastic errors are drawn from a AR(1) process.

**Table 2.18:** Information ratio: Treynor-Black using Student's-t

Stocks= 10; IR= 0.3214409							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.0370	0.0087	0.0235	0.0220	0.0340	0.0860
Bayesian NI		-0.2367	0.1923	0.2279	0.2070	0.2550	0.3150
Bayesian I	120	0.1998	0.2785	0.2905	0.2877	0.2995	0.3162
MELO NI		-0.1551	0.2033	0.2387	0.2233	0.2632	0.3135
MELO I		0.2064	0.2796	0.2916	0.2885	0.3005	0.3159
Naive (1/N)		0.0490	0.0490	0.0490	0.0490	0.0490	0.0490
Plug-in		-0.0480	0.0131	0.0295	0.0284	0.0429	0.0999
Bayesian NI		-0.1994	0.2386	0.2614	0.2554	0.2789	0.3113
Bayesian I	240	0.2449	0.2959	0.3029	0.3012	0.3081	0.3193
MELO NI		-0.1334	0.2469	0.2680	0.2612	0.2825	0.3114
MELO I		0.2498	0.2960	0.3027	0.3012	0.3084	0.3195
Naive (1/N)		0.0490	0.0490	0.0490	0.0490	0.0490	0.0490
Stocks= 50; IR=1.909063							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.0599	0.0157	0.0341	0.0298	0.0484	0.1003
Bayesian NI		0.7274	0.9536	1.0237	1.0199	1.0819	1.3111
Bayesian I	120	1.1116	1.3356	1.3978	1.3946	1.4559	1.6795
MELO NI		0.9739	1.2443	1.3206	1.3145	1.3889	1.6045
MELO I		1.1092	1.3346	1.3962	1.3930	1.4544	1.6618
Naive (1/N)		0.0882	0.0882	0.0882	0.0882	0.0882	0.0882
Plug-in		-0.0674	0.0336	0.0507	0.0456	0.0639	0.0997
Bayesian NI		1.0338	1.2289	1.2755	1.2728	1.3169	1.4602
Bayesian I	240	1.4132	1.5688	1.6103	1.6065	1.6452	1.7830
MELO NI		1.3343	1.5255	1.5672	1.5643	1.6082	1.7448
MELO I		1.4194	1.5685	1.6080	1.6053	1.6443	1.7888
Naive (1/N)		0.0882	0.0882	0.0882	0.0882	0.0882	0.0882

Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors. Stochastic errors are drawn from a multivariate student's-t distribution with 4 degrees of freedom.

**Table 2.19:** Information ratio: Treynor-Black using AR(1)

Stocks= 10; IR= 0.3214409							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.0499	0.0032	0.0172	0.0160	0.0292	0.0805
Bayesian NI		-0.1766	-0.0161	0.0636	0.0424	0.1024	0.1968
Bayesian I	120	-0.2033	0.1379	0.1579	0.1410	0.1763	0.2473
MELO NI		-0.1529	0.0306	0.0740	0.0643	0.1077	0.2046
MELO I		-0.1516	0.1293	0.1506	0.1387	0.1690	0.2473
Naive (1/N)		0.0490	0.0490	0.0490	0.0490	0.0490	0.0490
Plug-in		-0.0555	0.0105	0.0250	0.0236	0.0369	0.1297
Bayesian NI		-0.2004	0.0404	0.0973	0.0627	0.1265	0.2318
Bayesian I	240	-0.2007	0.1623	0.1771	0.1729	0.1917	0.2471
MELO NI		-0.1424	0.0550	0.1000	0.0815	0.1275	0.2239
MELO I		-0.1397	0.1572	0.1719	0.1690	0.1868	0.2400
Naive (1/N)		0.0490	0.0490	0.0490	0.0490	0.0490	0.0490
Stocks= 50; IR=1.909063							
Method	Sample size	Min	1st Q	Median	Mean	3rd Q	Max
Plug-in		-0.0552	0.0010	0.0221	0.0191	0.0400	0.0904
Bayesian NI		-0.1212	0.0337	0.0682	0.0599	0.0966	0.1776
Bayesian I	120	0.1136	0.1824	0.1984	0.1985	0.2156	0.2848
MELO NI		-0.0727	0.0477	0.0798	0.0727	0.1049	0.2142
MELO I		0.1017	0.1852	0.2007	0.2008	0.2171	0.2816
Naive (1/N)		0.0882	0.0882	0.0882	0.0882	0.0882	0.0882
Plug-in		-0.0662	0.0169	0.0353	0.0310	0.0514	0.0947
Bayesian NI		-0.1258	0.0800	0.1068	0.1034	0.1333	0.2247
Bayesian I	240	0.1742	0.2219	0.2366	0.2370	0.2510	0.3236
MELO NI		-0.0772	0.0907	0.1152	0.1118	0.1376	0.2275
MELO I		0.1776	0.2232	0.2388	0.2382	0.2520	0.3211
Naive (1/N)		0.0882	0.0882	0.0882	0.0882	0.0882	0.0882

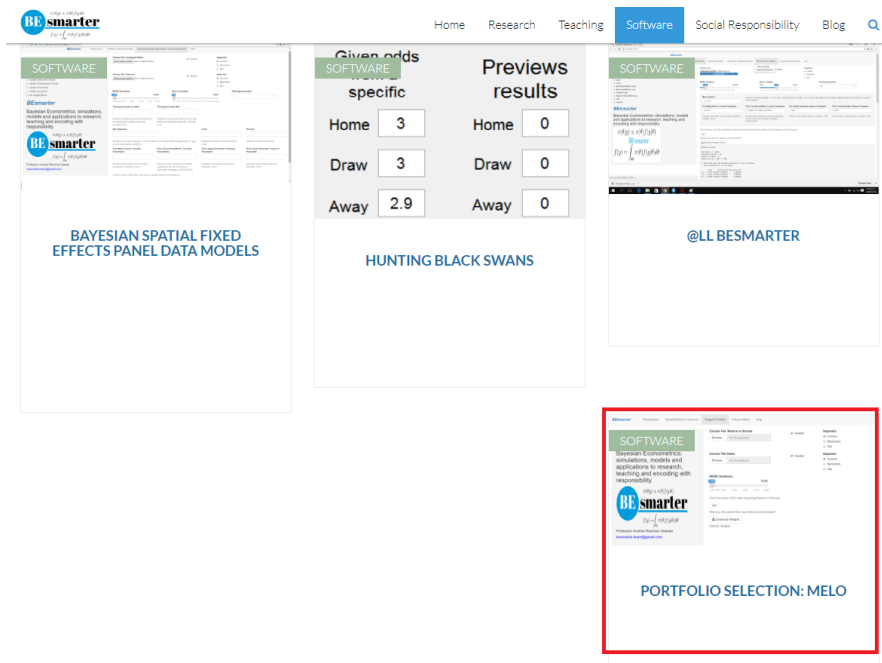
Note: This table shows summary statistics for the squared error.

“NI”: Non-informative priors. “I”: Informative priors.

Stochastic errors are drawn from a AR(1) process.

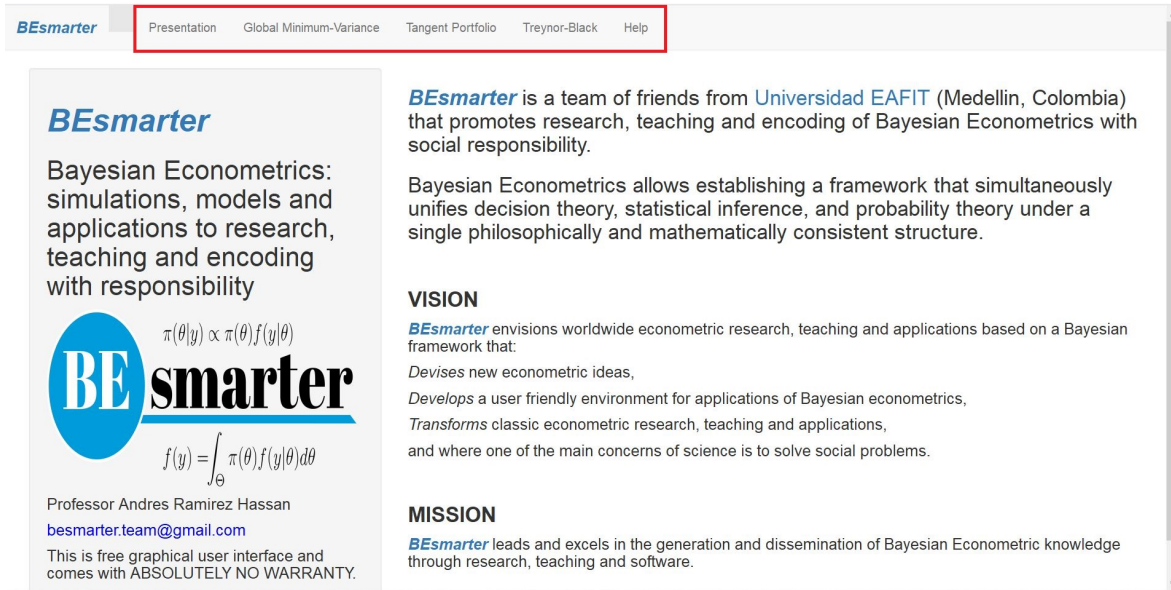
### 3 GUI user instructions

We have developed a graphical user interface (GUI) using Shiny to facilitate the implementation of the methodologies proposed in this paper. Here, we present a brief description of how it works. First, go to <http://www.besmarter-team.org/software>. This web page has other applications, so “PORTFOLIO SELECTION: MELO” should be selected (see Figure 3.1).



**Fig. 3.1:** GUI: BeSmarter software

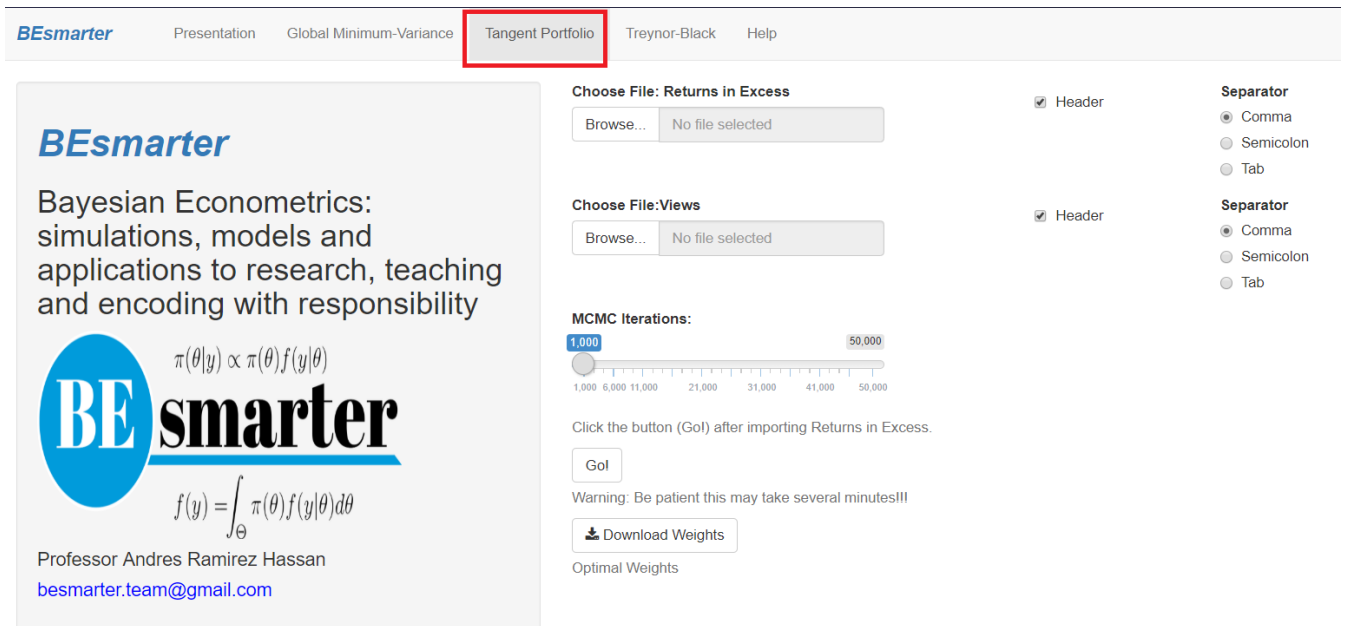
Figure 3.2 shows our GUI. It has five tap panels: Presentation, Global Minimum-Variance, Tangency Portfolio, Treynor–Blacka, and Help. The Presentation panel contains information about our research and the Help panel contains this document as a reference.



**Fig. 3.2:** GUI: Presentation

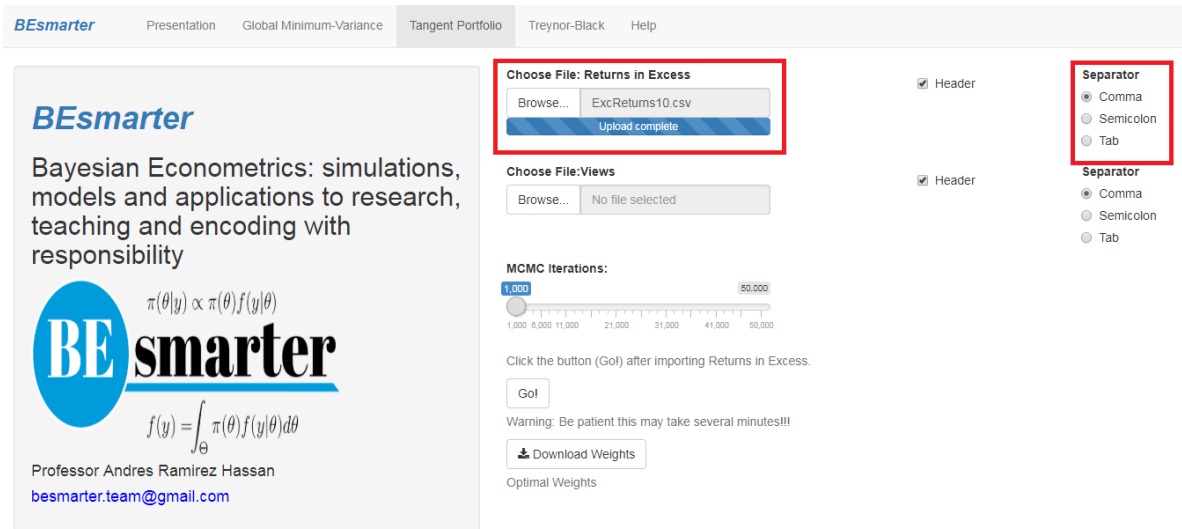
We describe how to use tangency portfolio and Treynor–Black panels, the panel for global minimum variance operates in a similar. Clicking on “Tangency Portfolio” leads to the window shown in Figure 3.3. To implement the tangency portfolio trade strategy, it is necessary to have a csv file with the excess of returns. In this file, each column must represent the historical excess of return for each asset in the portfolio. Additionally, the columns can be separated by comma, semicolon, or tab and must be specified in the right side of the panel (Figure 3.4). The last file should be selected from your computer by clicking the bottom “Choose File: Returns in Excess”. In case you want to use the informative cases, you should have a csv file with the information about the expected returns, which must be a column vector. The bottom “Choose File: Views” is used to load the views about the expected returns.





**Fig. 3.3:** GUI: Tangency portfolio

The window that is given once the excess of returns are loaded is shown in Figure 3.4. Then, clicking the bottom “Go” button calculates the portfolio weights.



**Fig. 3.4:** GUI: Excess of Returns loaded

The weights are shown in the window in Figure 3.5. Given that we did not load any file with the views, the weights for the methodologies that use informative priors are not calculated. However, Figure 3.6 shows the window if you also load a file with the views for the returns. Finally, if you want to download the weights in a csv file, just click the “Download Weights” bottom.

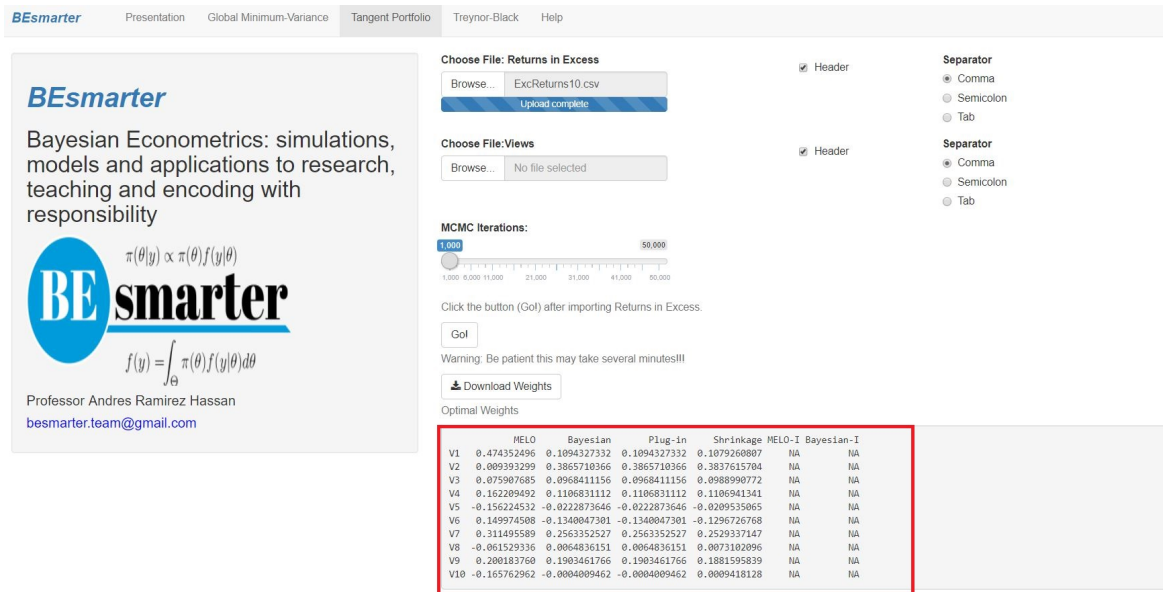


Fig. 3.5: GUI: Portfolios weights

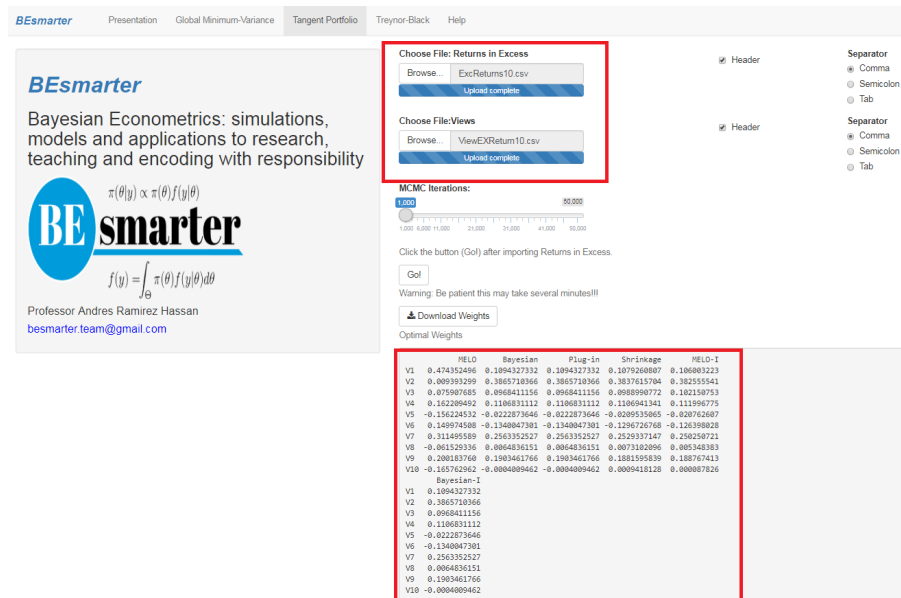


Fig. 3.6: GUI: Portfolios weights including informative methodologies

Now, if you want to apply the Treynor–Black trading strategy, you should click on the panel “Treynor–Black”, as shown in Figure 3.7. For the implementation of this methodology, we need not only the historical excess of returns for each asset in the portfolio but also the historical excess of returns of the factor. Thus, for the excess of return, a csv file (as described in the tangency portfolio panel) and a different csv file with the historical information for the excess of returns of the factor associated with the model are required.<sup>4</sup>

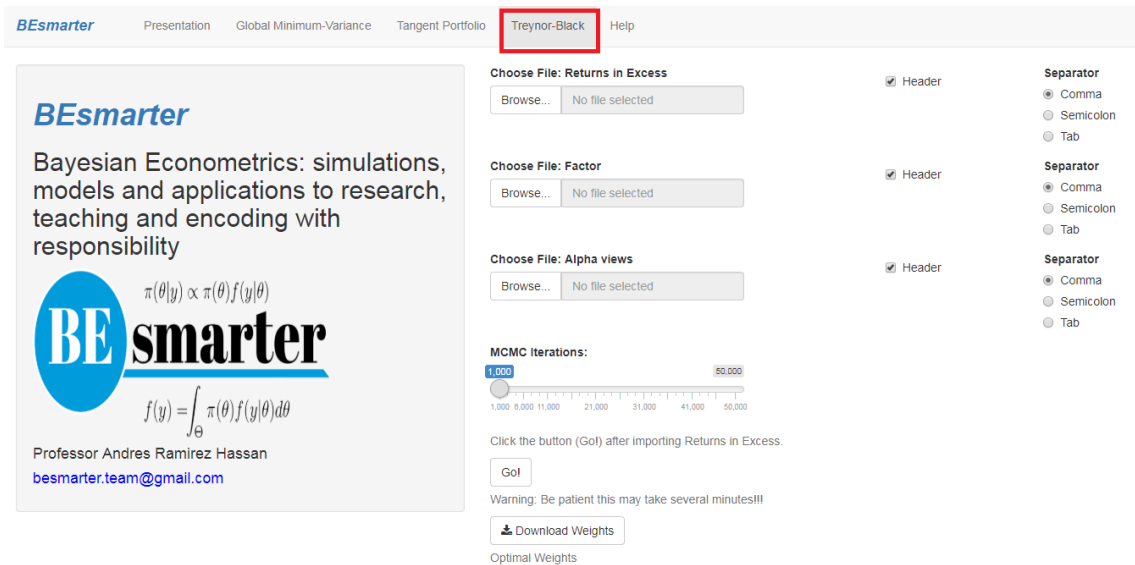


Fig. 3.7: Treynor–Black panel

The panel after loading the historical information for excess of returns of the assets and the factor is shown in Figure 3.8. Then, you just need to click the “GO” button to estimate the weights of the portfolio. The weights look like Figure 3.9. Finally, in case you have information about the abnormal returns of the assets, you could incorporate it by the bottom “Choose file: Alpha views”. This information must be a column vector saved in a csv file (Figure 3.10). Finally, a csv file with the weights can be obtained by clicking the “Download Weights” button at the bottom of the page.

<sup>4</sup> This must be a column vector.

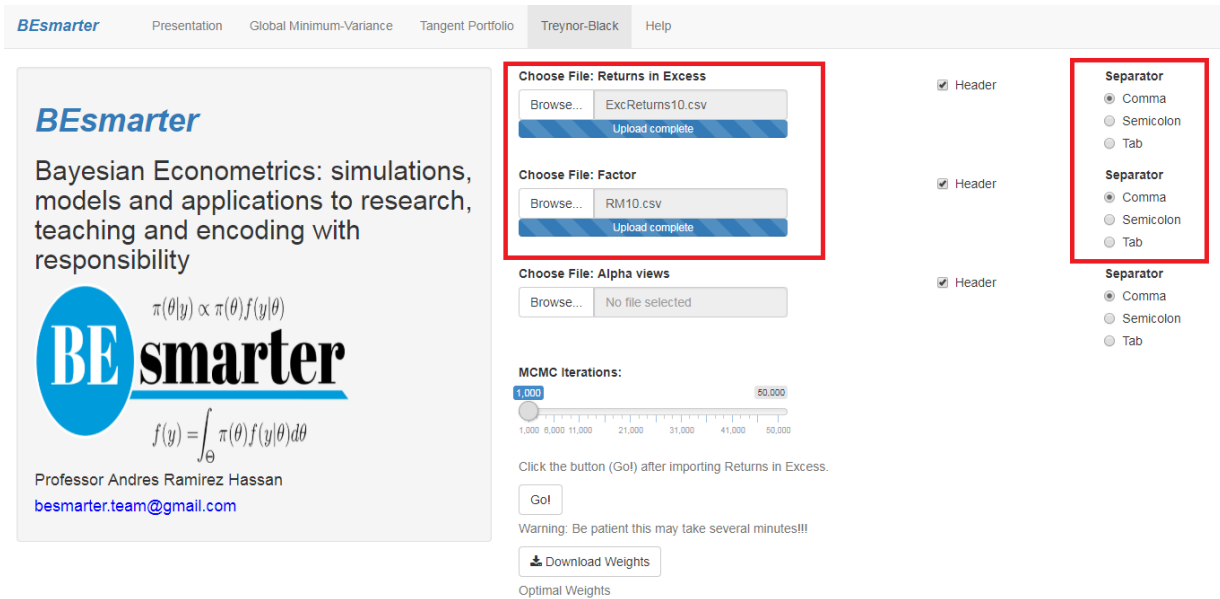


Fig. 3.8: Treynor-Black panel

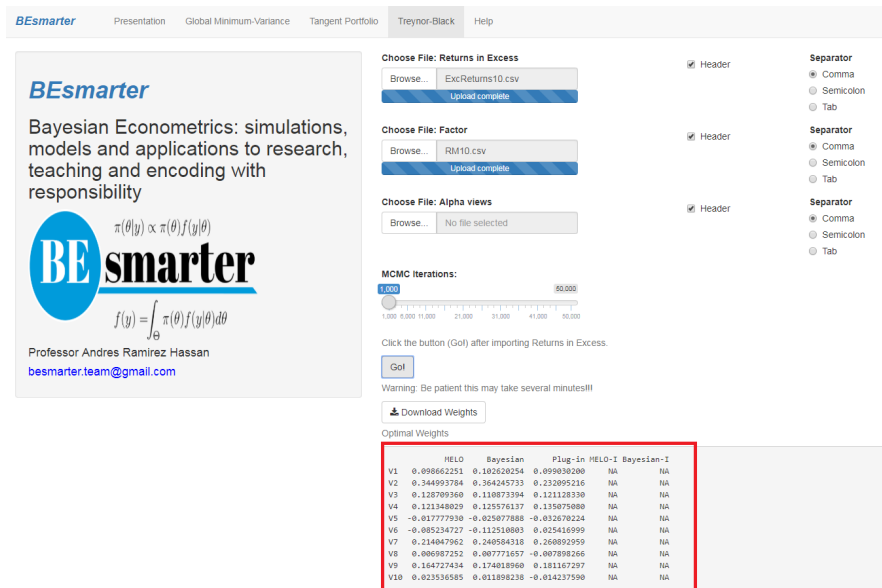



Fig. 3.9: Treynor-Black weights non-informative

BEsmarter Presentation Global Minimum-Variance Tangent Portfolio Treynor-Black Help

**BEsmarter**

Bayesian Econometrics: simulations, models and applications to research, teaching and encoding with responsibility

$\pi(\theta|y) \propto \pi(\theta)f(y|\theta)$ 

 $f(y) = \int_{\Theta} \pi(\theta)f(y|\theta)d\theta$

Professor Andres Ramirez Hassan  
besmarter.team@gmail.com

**Choose File: Returns in Excess**

Browse... ExcReturns10.csv  Header Separator  
 Comma  
 Semicolon  
 Tab

**Choose File: Factor**

Browse... RM10.csv  Header Separator  
 Comma  
 Semicolon  
 Tab

**Choose File: Alpha views**

Browse... ViewEXReturn10.csv  Header Separator  
 Comma  
 Semicolon  
 Tab

MCMC iterations: 50,000

Click the button (Go) after importing Returns in Excess.

Go

Warning: Be patient this may take several minutes!!!

**Download Weights**

Optimal Weights

	HELO	Bayesian	PJug-in	HELO-I	Bayesian-I
V1	0.09495177	0.102620254	0.099830200	0.09637399	0.104895126
V2	0.33837002	0.364245733	0.232895216	0.35845450	0.371700757
V3	0.12099994	0.110873394	0.121128330	0.11187649	0.106187635
V4	0.12593184	0.125576137	0.135075080	0.12206996	0.120002952
V5	-0.01902354	-0.025077888	-0.032670224	-0.01735996	-0.024146056
V6	-0.08111397	-0.112510803	0.025416999	-0.10426252	-0.119688208
V7	0.21977847	0.240584318	0.260892959	0.23576864	0.245843982
V8	0.01512158	0.00771657	-0.007692656	0.01054121	0.007341545
V9	0.16038431	0.174018960	0.181167297	0.17468070	0.179471060
V10	0.02359758	0.011898238	-0.014237590	0.01285701	0.007791207

Fig. 3.10: Treynor-Black weights with prior information

## References

1. A. Zellner, *An introduction to Bayesian inference in econometrics* (John Wiley, 1996)
2. E. Greenberg, *Introduction to Bayesian Econometrics*, vol. 85 (2009). DOI 10.1111/j.1475-4932.2009.00578.x. URL <http://www.amazon.de/dp/0521858712>
3. G.c. Tiao, A. Zellner, *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **26**(2), 277 (1964). URL <http://www.jstor.org/stable/2984424>
4. J.M. Dickey, *The Annals of Mathematical Statistics* **38**(2), 511 (1967). URL <http://www.jstor.org/stable/2239163>
5. E. Lehmann, G. Casella, *Theory of Point Estimation*, second edition edn. (Springer, 2003)
6. P.J. Bickel, J.A. Yahav, *Z. Wahrsch. Verw. Geb.* **11**, 257 (1969)
7. R. Kass, D. Steffey, *Journal of the American Statistical Association* **84**(407), 717 (1989)