1 Priors

The procedure consists of two steps:

1. Estimate the model for the training sample
2. Estimate the model for the remainder of the sample using the draws from the first step to form a prior

Both for Student t and normal errors we specify two prior distributions: for the coefficients $\gamma \sim \mathcal{N}_p\left(\boldsymbol{\gamma}|\boldsymbol{\gamma_0}, G_0\right)$ and for the precision matrix $\Omega^{-1} \sim \mathcal{W}_D\left(\Omega^{-1}|\rho_0, R_0\right)$. Parameters for the two distributions are specified differently for each step.

1.1 Step 1

Regression coefficients:

$$\begin{aligned}\gamma_0 &= 0 \\ G_0 &= C_1^2 I_p\end{aligned}$$

The prior distribution of the coefficient vector is centered around zero. The covariance matrix is assumed to be diagonal. We use the value $C_1 = 2$.
Precision matrix:

$$\begin{aligned}\rho_0 &= d + 8 \\ R_0 &= \frac{1}{C_2^2(\rho_0 - d - 1)} I_d\end{aligned}$$

The mean value of Ω implied by the prior is a diagonal matrix $C_2^2 I_d$. We set $C_2 = 0.05$.

1.2 Step 2

Here we use the draws based on the training sample to construct priors. Denote posterior means of draws of *gamma* as $\overline{\gamma}$ and the sample covariance matrix as $\widehat{G}$. We also calculate a posterior mean of the Ω^{-1} draws: $\overline{\Omega}^{-1}$.
Regression coefficients :

$$\begin{aligned}\gamma_0 &= \overline{\gamma} \\ G_0 &= C_3^2 \widehat{G}\end{aligned}$$

The prior is centered around the posterior mean of the first step draws. The standard deviation is based on the sample standard deviation of the first stage draws adjusted by the factor of C_3 to reflect uncertainty. We set $C_3 = 3$.
Precision matrix:

$$\begin{aligned} \rho_0 &= d + C_4 \\ R_0 &= \frac{1}{\rho - d - 1}\overline{\Omega}^{-1} \end{aligned}$$

The prior is constructed to set the mean equal to the posterior mean of Ω^{-1} based on the training sample. One way of widening the prior would be to decrease number of degrees of freedom ρ_0 by adjusting the value of C_4. Currently we use $C_4 = 6$.

2 Motivational Simulation

In order to motivate our research we simulate asset returns and demonstrate that the true model is selected. The returns are assumed to be follow the famous Fama-French 3 factor(Mkt.RF, HML and SMB) structure without intercept. The standard errors follow Student-t distribution with 4 degrees of freedom. The simulation is based on posterior means obtained when fitting the model to the real data. Other parameters are the same as in the original sample. The pool of candidate models includes all combinations of Fama-French 5 factors(Mkt.RF, HML, SMB, RMW and CMA) and a constant. Considered distributions include normal and Student-t with 3, 4, 6, 8 and 12 degrees of freedom. We fit in total $6 \times 2^6 = 384$ model.
The simulation setup is described below:

1. Fit the Fama-French 3 factor model without an intercept to 10 value-weighted industry portfolios using the full sample. The errors are assumed to follow Student-t distribution with 4 degrees of freedom.

2. Simulate a dataset assuming the true parameters γ and Ω^{-1} to be equal to the posterior means:

    - Simulate errors:

    $$\boldsymbol{\varepsilon}_t^s \sim t_{10,4}\left(0, \Omega\right)$$

    - Simulate returns using values of Fama-French 3 factors observed in the data:

    $$\mathbf{y}_t^s = X_t \boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t^s$$

Factor	Student_3	Student_4	Student_6	Student_{12}	Normal
constant					
Mkt.RF	✓	✓	✓	✓	✓
SMB	✓	✓	✓	✓	✓
HML	✓	✓	✓	✓	✓
RMW					
CMA					
log Likelihood	10027.7	10361.34	10276.61	10028.09	10102.35

3. Estimate all candidate models and evaluate the likelihood. Run the usual two step estimation procedure using the training sample to construct priors for each model.

As can be seen from the table above that lists the factors of the best model for each distribution, the true model is always selected.

This table show the best three models based on Student-t errors with 4 degrees of freedom (the true distribution).

Factor	Model 1	Model 2	Model 3
constant			✓
Mkt.RF	✓	✓	✓
SMB	✓	✓	✓
HML	✓	✓	✓
RMW		✓	
CMA			
log Likelihood	10027.7	10264.24	10263.19

Example of a table for the big run:

Factor	Student_4	Student_{12}	Normal
constant			
LIQv			
MOM	✓	✓	✓
Mkt.RF	✓	✓	✓
SMB			✓
HML	✓	✓	✓
RMW	✓	✓	✓
CMA	✓	✓	✓
QMJ	✓	✓	✓
ME	✓	✓	✓
IA			
ROE	✓	✓	✓
HMLDev			✓
log Likelihood	9684.65	9669.05	9517.35