On the Long Run Volatility of Stocks

November 15, 2017

Appendix 1: Predictive Variance Decomposition

Let $\Theta = (\alpha, \beta, \Sigma)$ represent all the fixed parameters in the model defined in (1). Using basic properties of expectations and variances, the predictive variance in (3), can be written as:

$$\begin{aligned}\operatorname{Var}\left(r_{T,T+k}|D_T\right) &= \mathrm{E}\left\{\operatorname{Var}(r_{T,T+k}|\mu_T, \Theta, D_T)\right\} + \operatorname{Var}\left\{\mathrm{E}(r_{T,T+k}|\mu_T, \Theta, D_T)\right\} \quad (1)\\ &= \mathrm{E}\left\{\operatorname{Var}(r_{T,T+k}|\mu_T, \Theta, D_T)\right\} \\ &+ \mathrm{E}\left\{\operatorname{Var}(\mathrm{E}(r_{T,T+k}|\mu_T, \Theta, D_T)\right\} + \operatorname{Var}\left\{\mathrm{E}\left[\mathrm{E}(r_{T,T+k}|\mu_T, \Theta, D_T)\right]\right\}. \quad (2)\end{aligned}$$

where μ_T is the last expected return, at time T.

The first term of the right hand side can now be expanded by using the properties of a MA(∞) process as described in the Appendix section of **?** and it takes the following form:

$$\operatorname{Var}\left(r_{T,T+k}|\mu_T, \Theta, D_T\right) = k\sigma_u^2\left[1 + 2\bar{d}\rho_{uw}A(k) + \bar{d}^2B(k)\right] \quad (3)$$

where

$$\begin{aligned}A(k) &= 1 + \frac{1}{k}\left(-1 - \beta\frac{1-\beta^{k-1}}{1-\beta}\right), \quad (4)\\ B(k) &= 1 + \frac{1}{k}\left(-1 - 2\beta\frac{1-\beta^{k-1}}{1-\beta} + \beta^2\frac{1-\beta^{2(k-1)}}{1-\beta^2}\right), \quad (5)\\ \bar{d}^2 &= \frac{1+\beta}{1-\beta}\frac{R^2}{1-R^2}, \quad (6)\\ R^2 &= \frac{\sigma_w^2}{\sigma_u^2(1-\beta^2)+\sigma_w^2}. \quad (7)\end{aligned}$$

The remaining terms follow directly from the forecast function of a state-space model and depend upon the posterior mean (m_T) and variance (C_T) of μ_T given D_T, and can be written as:

$$\begin{aligned}
\mathrm{Var}_{\mu_T}\left\{\mathrm{E}(r_{T,T+k}|\mu_T,\Theta,D_T)\right\} &= \left(\frac{1-\beta^k}{1-\beta}\right)^2 C_T, &(8)\\
\mathrm{E}_{\mu_T}\left[\mathrm{E}(r_{T,T+k}|\mu_T,\Theta,D_T)\right] &= k\frac{\alpha}{1-\beta}+\frac{1-\beta^k}{1-\beta}\left(m_T-\frac{\alpha}{1-\beta}\right). &(9)
\end{aligned}$$

Therefore we can decompose the predictive variance into 5 interpretable components following the nomenclature of **?**:

$$\begin{aligned}
\mathrm{Var}\left(r_{T,T+k}|D_T\right) = \mathrm{E}\left\{k\sigma_u^2|D_T\right\} &\quad \text{(i.i.d uncertainty)} &(10)\\
+\mathrm{E}\left\{k\sigma_u^2 2\bar{d}\rho_{uw}A(k)|D_T\right\} &\quad \text{(mean reversion)} &(11)\\
+\mathrm{E}\left\{k\sigma_u^2\bar{d}^2 B(k)|D_T\right\} &\quad \text{(future } \mu \text{ uncertainty)} &(12)\\
+\mathrm{E}\left\{\left(\frac{1-\beta^k}{1-\beta}\right)^2 C_T|D_T\right\} &\quad \text{(current } \mu \text{ uncertainty)} &(13)\\
+\mathrm{Var}\left\{k\frac{\alpha}{1-\beta}+\frac{1-\beta^k}{1-\beta}\left(m_T-\frac{\alpha}{1-\beta}\right)|D_T\right\} &\quad \text{(estimation risk)} &(14)
\end{aligned}$$

Each of the terms are evaluated via Monte Carlo taking as inputs the values of m_T, C_T and Θ in each step of the MCMC used for model fitting.

Appendix 2: Time-Varying Σ_t: Posterior Sampling Details

From Section 2.2 , we have the observation equations:

$$(u_t,w_t)' \sim N(0,\Sigma(\theta_t)),\quad \theta_t=(\theta_{t,1},\theta_{t,2},\theta_{t,3})$$

$$\begin{aligned}
u_t &= \exp\left(\frac{\theta_{t,1}}{2}\right) Z_{t1}\\
w_t &= \theta_{t,3}\,u_t+\exp\left(\frac{\theta_{t,2}}{2}\right) Z_{t2},
\end{aligned}$$

state equation

$$p(\theta_t \,|\theta_{t-1}) \propto q(\theta_t \,|\, \theta_{t-1}) \; f(\theta_t), \tag{15}$$

and initial state prior

$$p(\theta_0) \propto q(\theta_0) f(\theta_0).$$

As discussed in Section 2.2 , $q(\theta_t \,|\, \theta_{t-1})$ and $q(\theta_0)$ represent a "standard" specification and the function $f(\theta_t)$ is used to inject prior beliefs about θ_t.

While this approach makes the prior specification relatively straightforward, it complicates posterior draws since evaluation of the transition distribution involves the normalizing constant in 15. Recall that the AR(1) parameters in $q(\theta_{t,i} \,|\, \theta_{t-1,i})$ are denoted by (a_i, b_i, c_i). Let $(a, b, c) = \{(a_i, b_i, c_i), i = 1, 2, 3\}$. Making (a, b, c) explicit in the above, we have:

$$\begin{aligned} p(\theta_t \,|\theta_{t-1}, a, b, c) &\propto q(\theta_t \,|\, \theta_{t-1}, a, b, c) \; f(\theta_t) \\ &= q(\theta_t \,|\, \theta_{t-1}, a, b, c) \; f(\theta_t) \; K(\theta_{t-1}, a, b, c) \end{aligned}$$

where K denotes the normalizing constant. This normalizing constant will be present in draws of the states $\{\theta_t\}$ given (a, b, c) and draws of (a, b, c) given the states using the usual Gibbs sampler approach conditional on everything else.

Our approach is to discretize each of the three states so that $\theta_{ti} \in G_i = \{g_{i1}, g_{i2}, \ldots, g_{in_i}\}$, $i = 1, 2, 3$, giving a three dimensional grid of possible θ_t vectors. While three dimensional grids are large, we have found that be carefully keeping track of what has been already computed and using parallel computation of K when needed (using openmp in C++) we can get draws in a reasonable amount of time. For each i, we draw the sequence $\{\theta_{ti}\}$ conditional on the other two state sequences and (a, b, c) using the forward filtering backward sampling algorithm [**??**]. Notice this is possible because the discretization of the state space enables us to do the forward filtering and backward sampling exactly. Finally, we use a random walk Metropolis to draw (a, b, c) using joint proposals for (a_i, b_i, c_i).

Appendix 3: Time-Varying Σ_t: Simulation Study

To demonstrate the effectiveness of our prior specification strategy for $\rho_{t,uw}$ we run a simple simulation study where simulated data was generated from the model:

$$\begin{pmatrix} u_{t+1} \\ w_{t+1} \end{pmatrix} \sim N(0, \mathbf{\Sigma}_t),$$

where the path of $\mathbf{\Sigma}_t$ was chosen to mimic the time-varying relationship between annual returns in the U.S. and U.K. markets. This was achieved by first computing the posterior for the time-varying covariance matrix of U.S. and U.K. returns under the Cholesky model in Section 2 and using the posterior mean to generate the time series for (u, w). This data is such that the path of $\rho_{t,uw}$ tends to be above 0.5.

We the proceed to fit two models for this data: (i) the Cholesky model with a very flat prior for $\rho_{t,uw}$ and (ii) the Cholesky model that uses the prior strategy described in Section 2.2, with $\bar{\rho} = 0.7$ and $\kappa = 0.5$. Draws from the priors and posteriors of $\rho_{t,12}$, $\sigma_{t,1}$ and $\sigma_{t,2}$ for both models are presented in Figure 1 and 2. It is clear that using the information that the correlation is above 0.5 leads to more precise estimates of the true path of $\rho_{t,12}$. It also has an impact on how precise variances are estimated.

In summary, in the presence of relevant available information, the strategy proposed in Section 2.2 is an effective vehicle to bring the information to the analysis and generate more precise estimates of the quantities of interest.

Appendix 4: Monthly and Quarterly Data

A number of previous studies that address the long term volatility of stocks do so using data on different frequencies, either monthly or quarterly. For completeness of our presentation we re-run the models in Section 3 for the U.S. returns on a monthly and quarterly basis. We use the same prior specifications (adjusted for the change in time scale) and the results for the predictive standard deviation are presented in Figure 3. In these frequencies, we achieve a very different conclusion as the predictive variances tend to go up with the horizon. Looking at the monthly results (left panel) we see a very clear agreement between all models and the unconditional, model-free line. The quarterly results are a little different but all lines have a similar shape with the time-varying $\mathbf{\Sigma}_t$ line starting at a lower level as the return volatility in recent quarters of the dataset is lower than the stationary level.

Figures 4 and 5 show priors and posteriors in for β (left panel) and ρ (right panel) for the "weak" (top row) and "strong" (bottom row) prior settings. Figure 4 (monthly) shows that ρ_{uw} is inferred to be not too negative while β very low. For the quarterly data in Figure 5, both quantities start to look more like the result we see in the annual analysis: ρ_{uw} is inferred to be more negative and β points to a more persistent dynamic behavior for μ.

These analyses show a nice progression from the monthly to quarterly to annual results and indicates that in shorter frequencies the signal about the dependence structure in μ_t gets smaller relative to the noise of the system. This overwhelms our ability to filter out the dynamics of expected returns even in the presence of strong beliefs about ρ_{uw}.

In order to estimate the predictive variance of stocks in the long run, we need to be able to learn about the dynamic behavior of expected returns. Only through the structure present in the annual data we are able to get estimates of ρ_{uw} and β in ranges that imply a negative sloping predictive variance as a function of the horizon.

Appendix 5: Posterior Predictive Checks for International Data

Figures 6 to 13 present the posterior predictive checks for the iid, "weak" and "strong" prior settings in the U.K., Australia, Canada, Hong Kong, Europe, France, Germany and Japan.

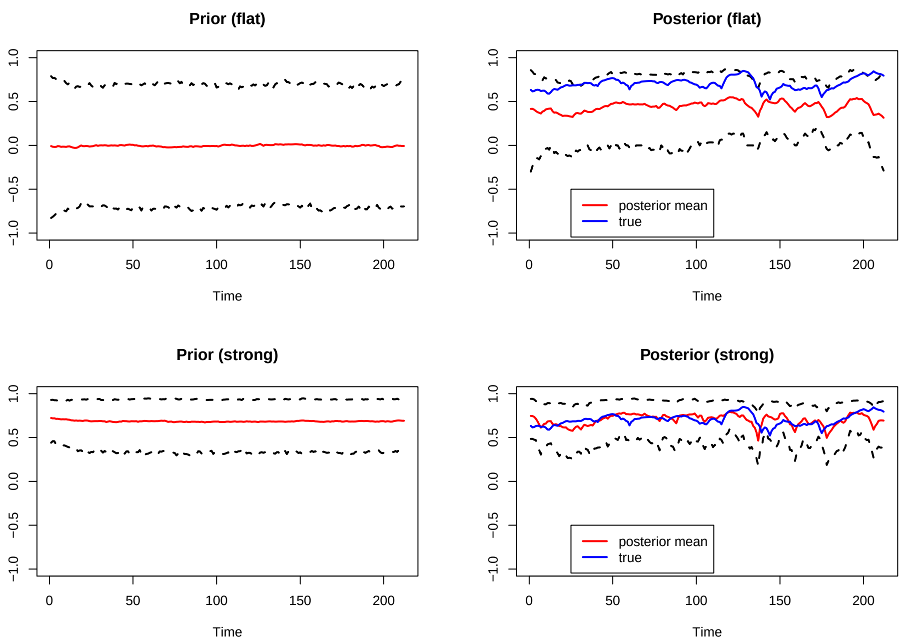

Figure 1: Priors (left) and posteriors (right) for $\rho_{t,uw}$ under a flat (top row) and a strong prior (bottom row). The solid red line in each plot represent the mean of each quantity while the two dashed lines give the 2.5% and 97.5 % quantiles of the distributions. The blue lines are the true path of correlations.

References

C. K. Carter and R. Kohn. On gibbs sampling for state space models. *Biometrika*, 81(3):541–553, 1994.

S. Frühwirth-Schnatter. Data augmentation and dynamic linear models. *Journal of time series analysis*, 15(2):183–202, 1994.

L. Pástor and R. F. Stambaugh. Are stocks really less volatile in the long run? *The Journal of Finance*, 67(2):431–478, 2012.

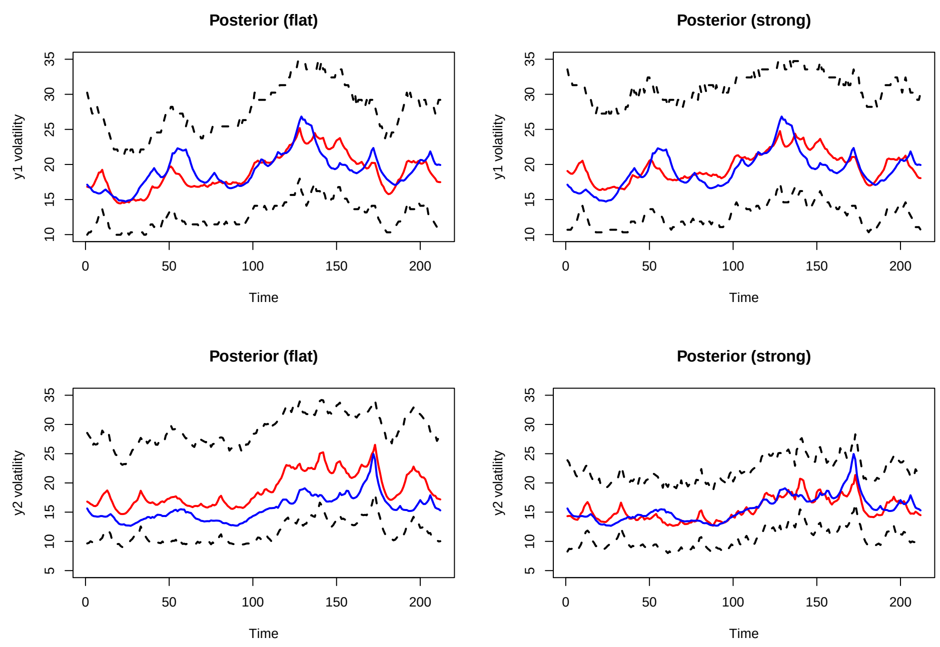

Figure 2: Posteriors (right) for σ_u (top row) and σ_w (bottom row) under a flat (left panels) and a strong prior (right row). The solid red line in each plot represent the mean of each quantity while the two dashed lines give the 2.5% and 97.5 % quantiles of the distributions. The blue lines are the true path of correlations

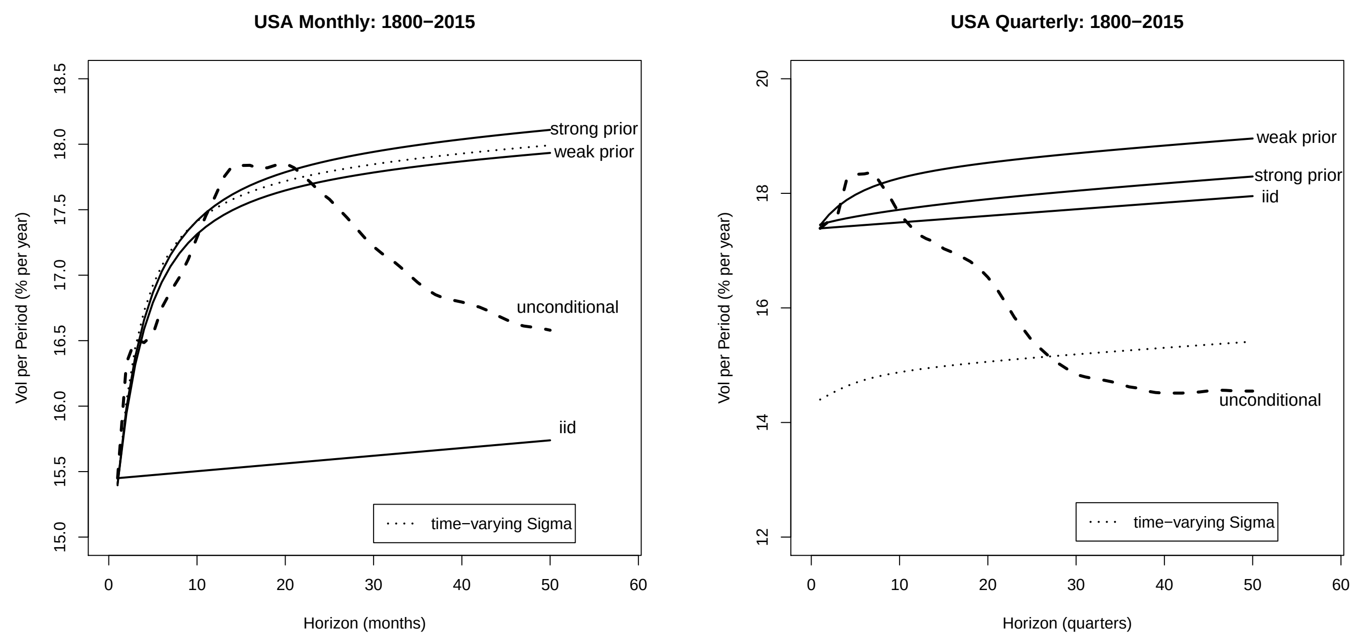

Figure 3: **U.S. monthly and quarterly returns:** Predictive volatility per period plotted for different horizons, i.e., $\sqrt{\frac{\mathrm{var}(r_{T,T+k}|D_T)}{k}}$.

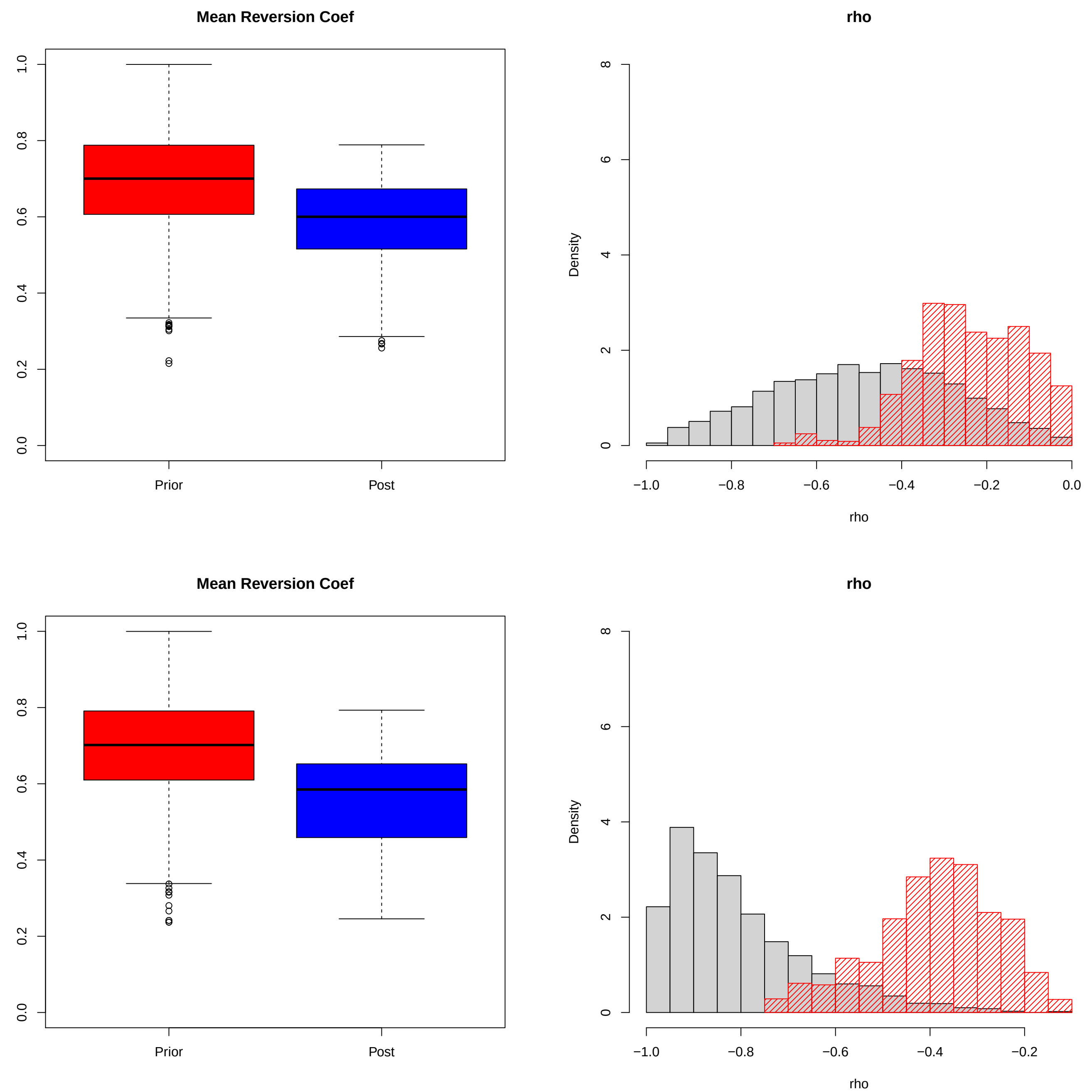

Figure 4: **U.S. monthly returns:** Top row: results for "weak prior". Bottom row: results for the "strong prior". Left panels show boxplots for priors (red) and posterior (blue) for β. Right panel shows prior (gray) and posterior (red) draws for ρ_{uw}.

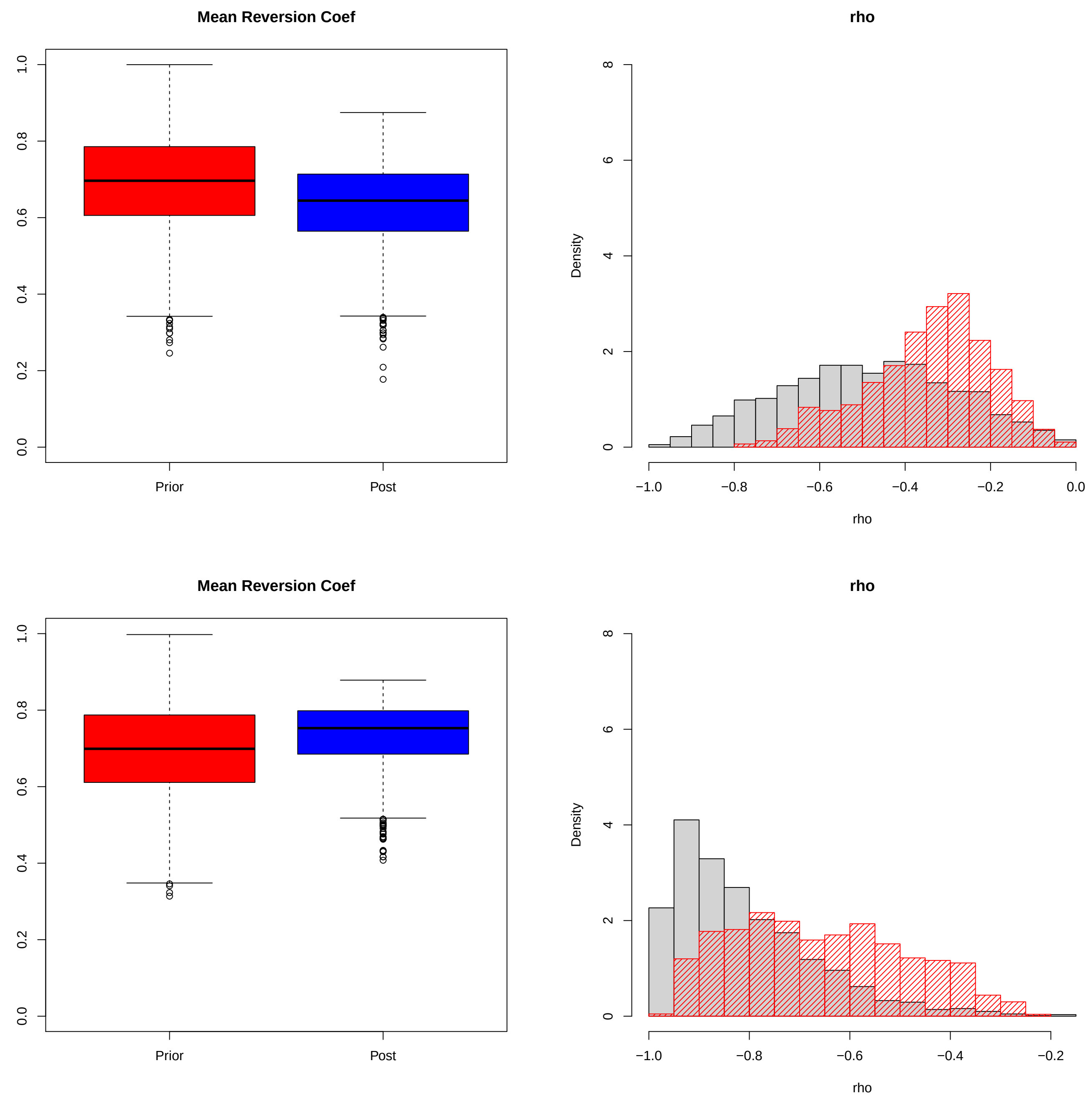

Figure 5: **U.S. quarterly returns:** Top row: results for "weak prior". Bottom row: results for the "strong prior". Left panels show boxplots for priors (red) and posterior (blue) for β. Right panel shows prior (gray) and posterior (red) draws for ρ_{uw}.

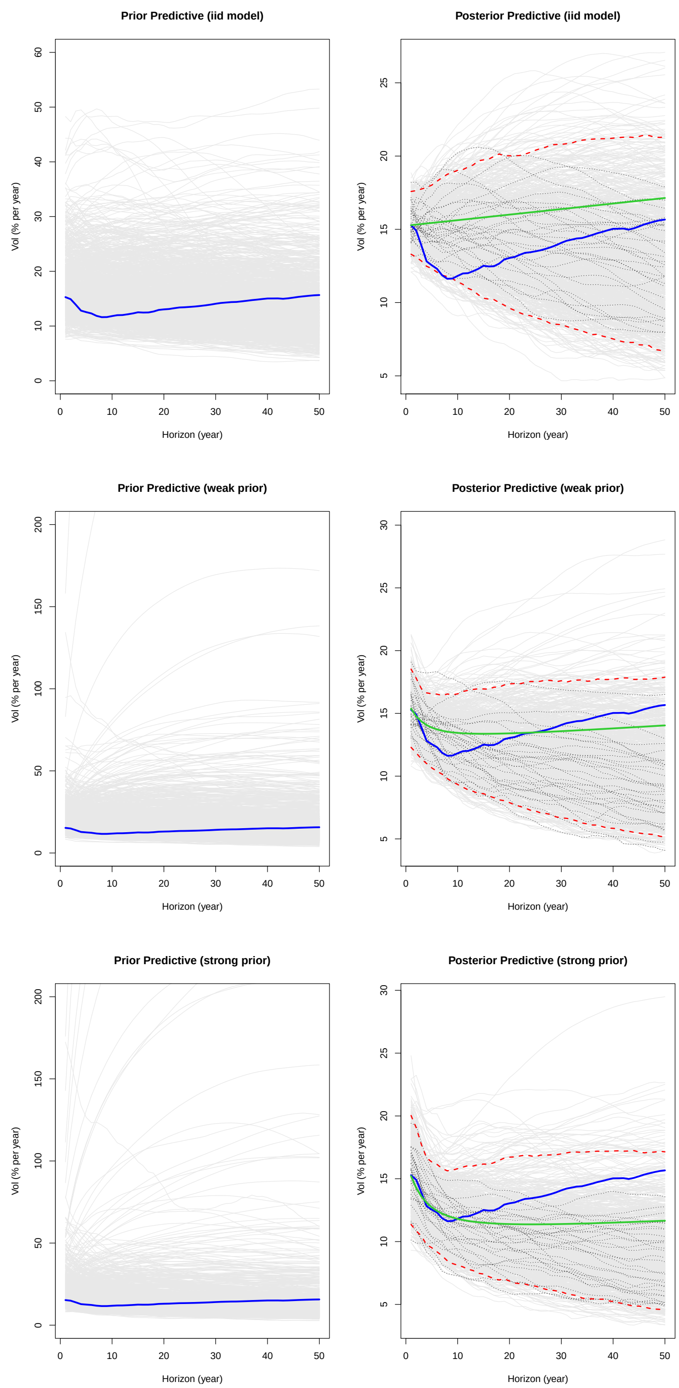

Figure 6: U.K. annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

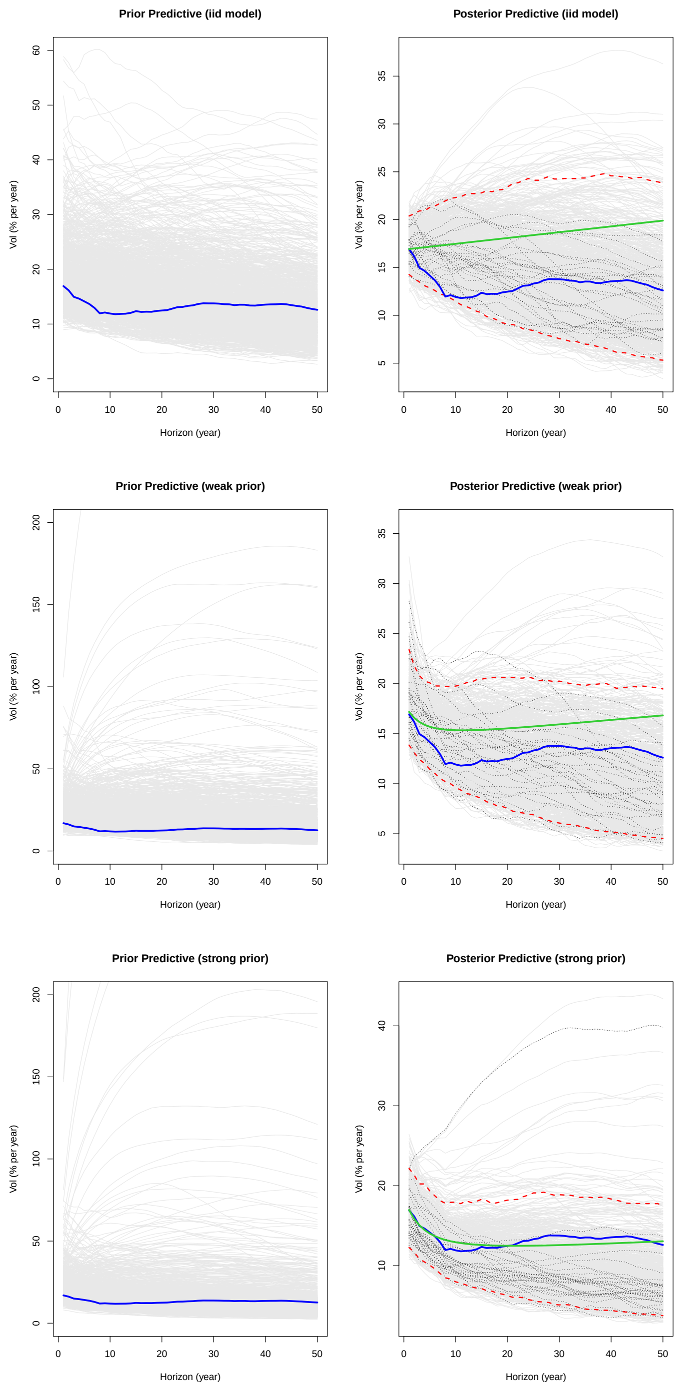

Figure 7: AUS annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), “weak” (middle row) and “strong” prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the “unconditional”, model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

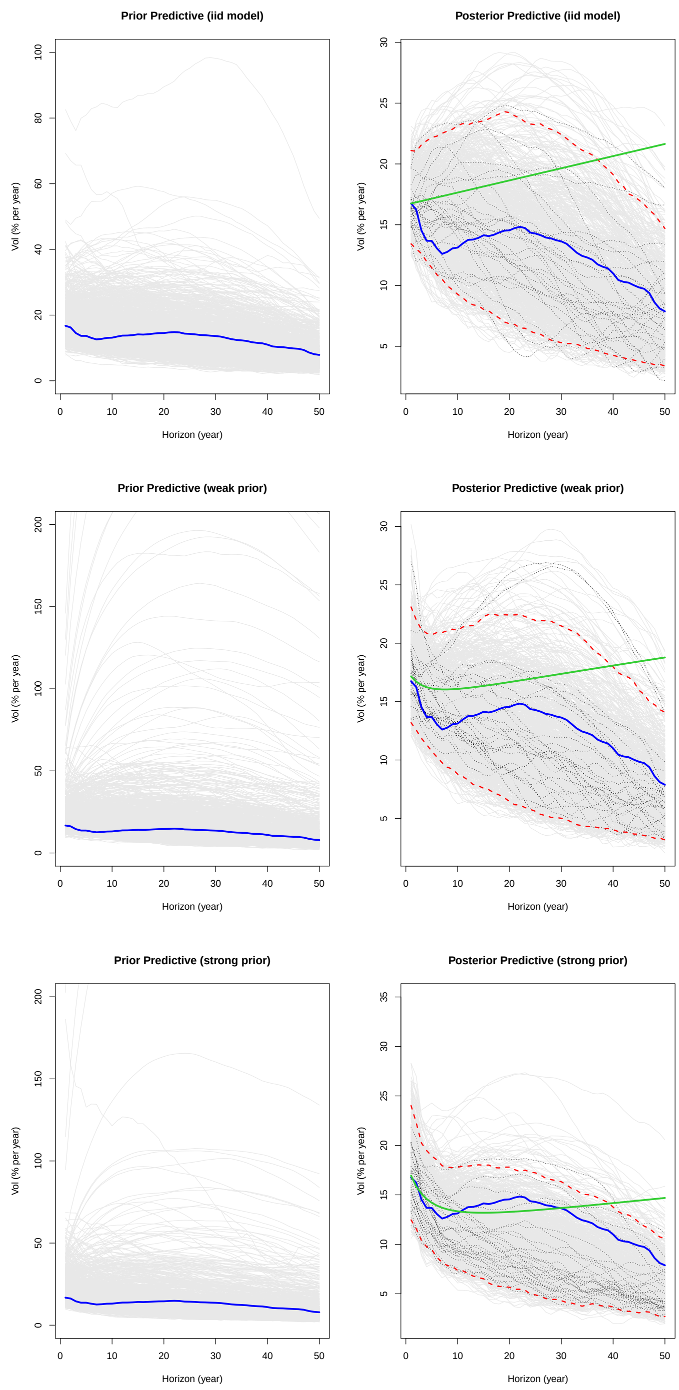

Figure 8: CAN annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

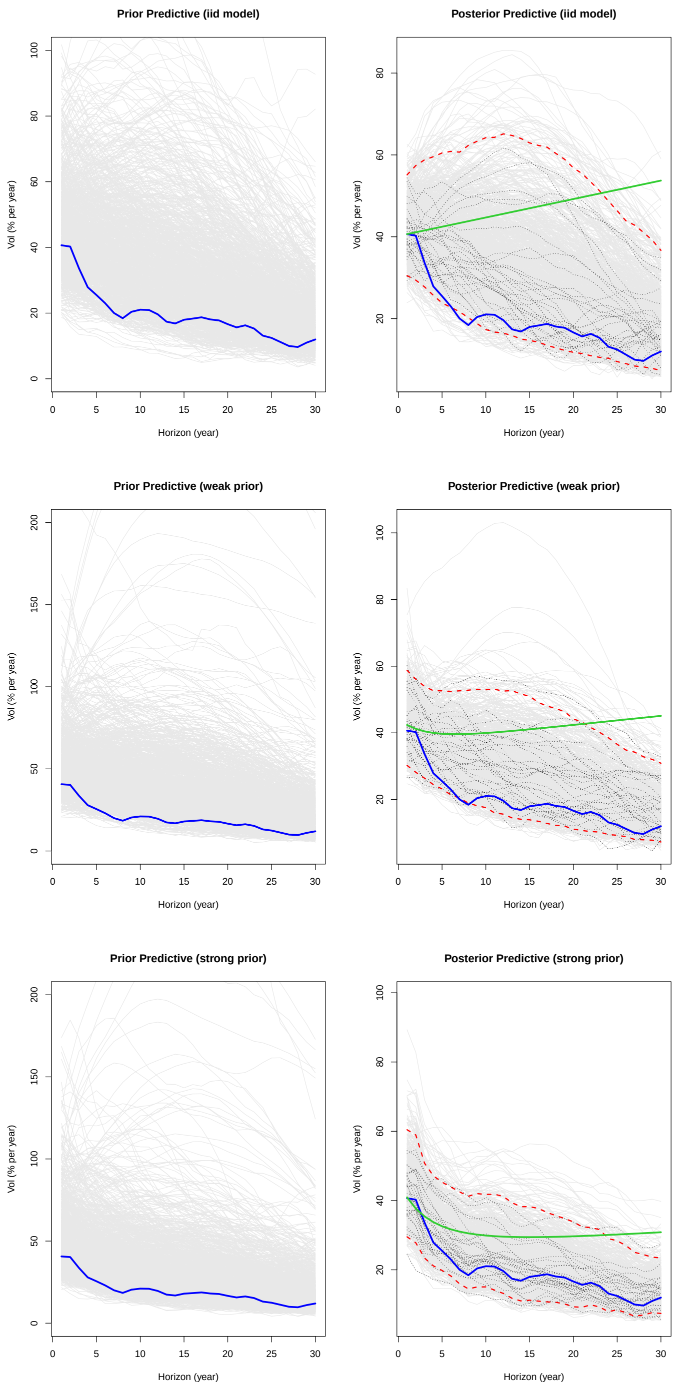

Figure 9: HK annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

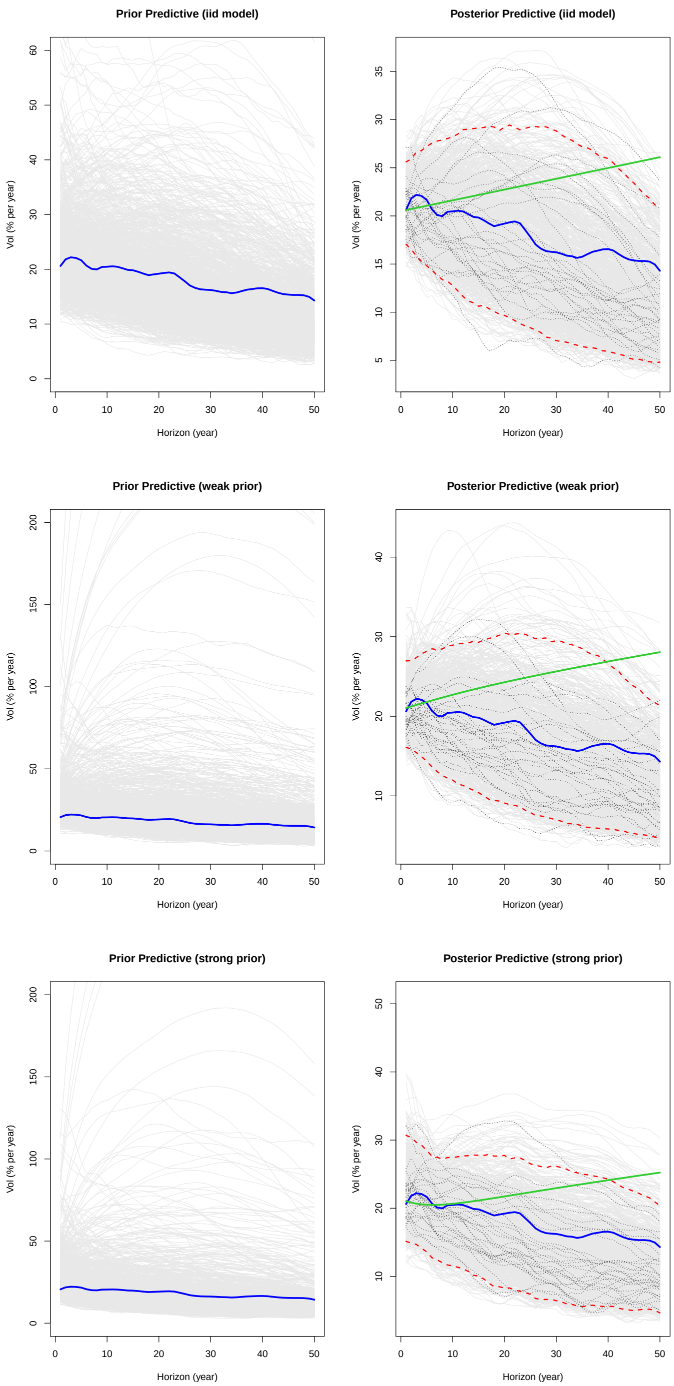

Figure 10: Europe annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

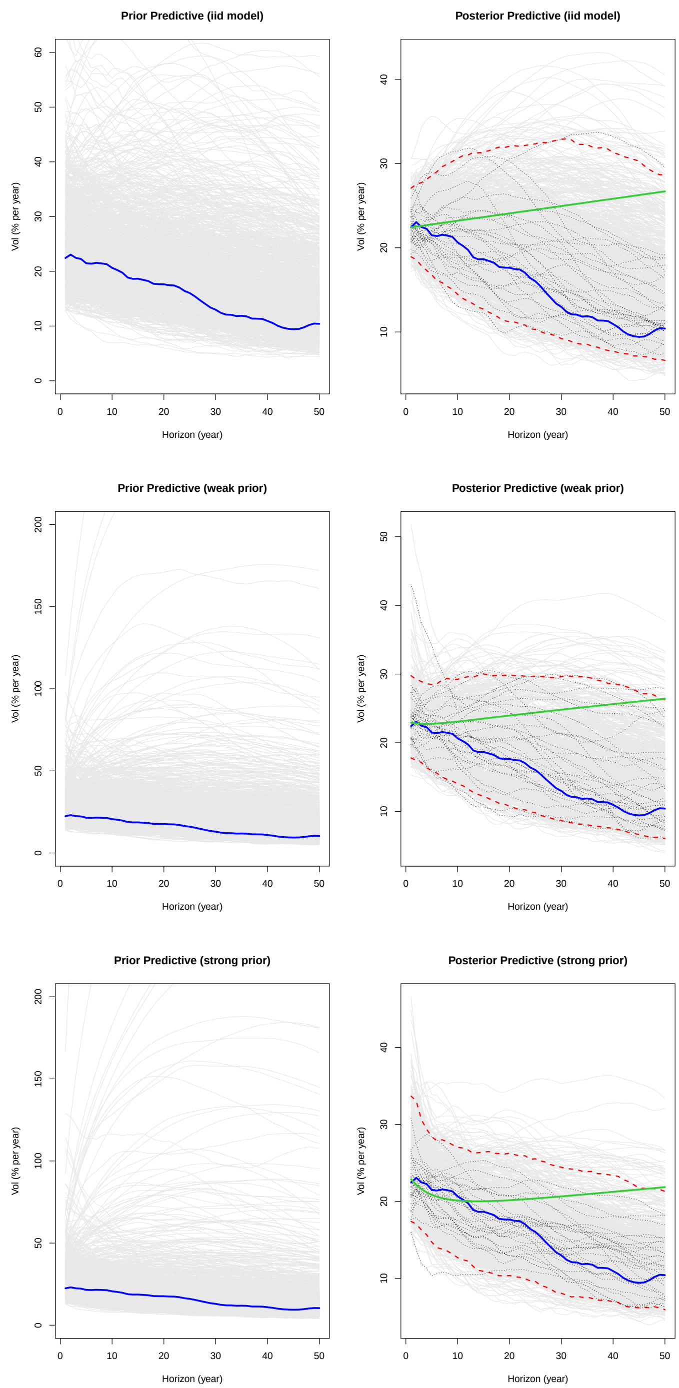

Figure 11: FRA annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

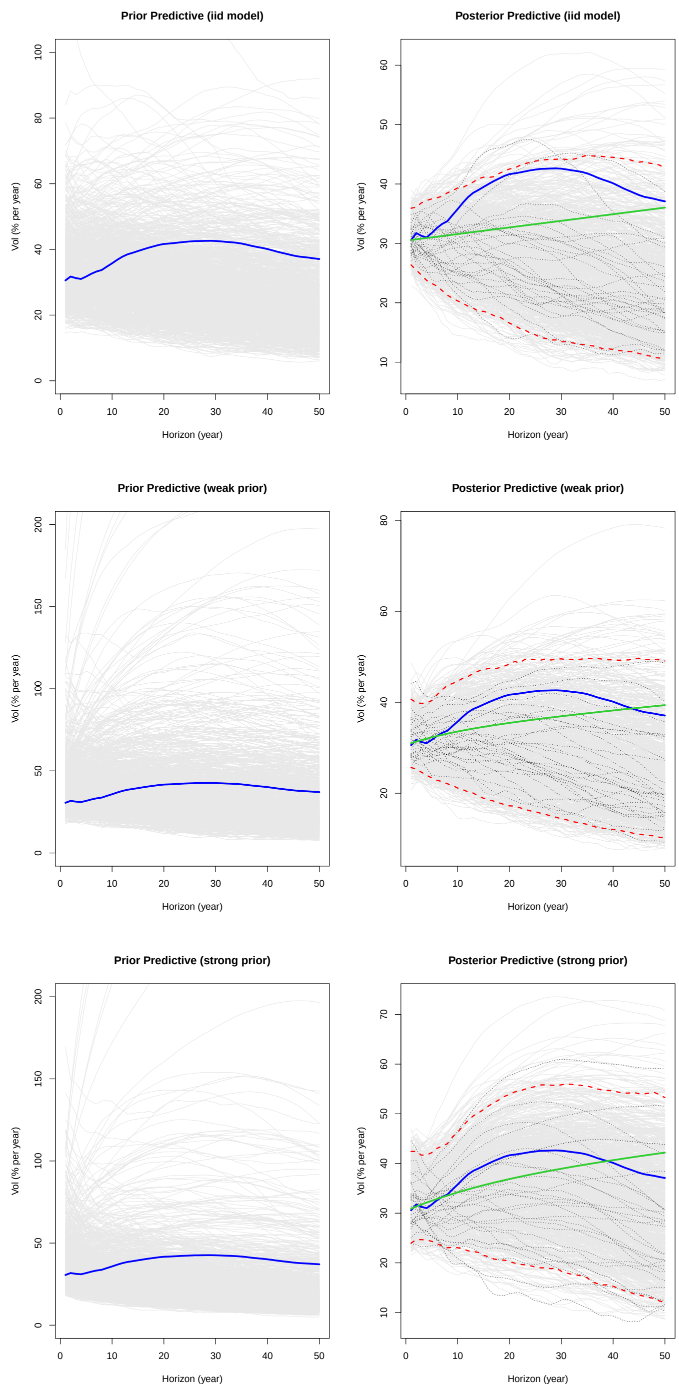

Figure 12: GER annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.

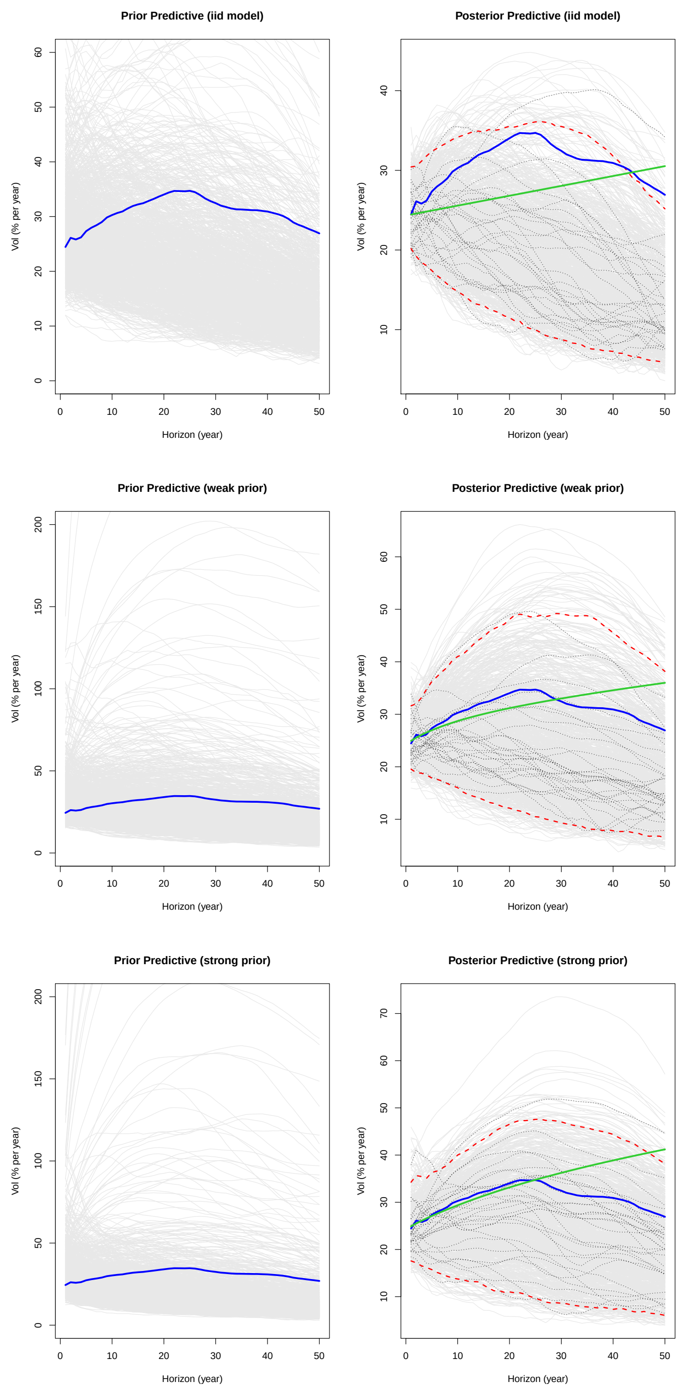

Figure 13: JPN annual returns. Prior (left panels) and posteriors (right panels) predictive simulations for the iid (top row), "weak" (middle row) and "strong" prior settings (bottom row). Gray lines are all the simulated draws. The blue line represents the statistic from the data, i.e., the "unconditional", model-free computation of the volatility per period. The black lines in the right panels are a subset of the gray draws plotted here to emphasize their shape. Red dotted lines represent the 2.5% and 97.5% quantiles of the gray draws. Green line is the predictive standard deviation per period resulting from the model considered in the plot.