Problem Set # 2

1. Identify the model parameters and propose a prior distribution for them.

   We consider the dynamic version of the Nelson-Siegel model proposed by Diebold et al. (2006), where the yield curve is driven by time-varying latent factors $(\beta_{0t}, \beta_{1t}, \beta_{2t})$. The parameters to be estimated are:

   - $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)$: unconditional means of the factors;
   - $\mathbf{A}$: a 3×3 autoregressive matrix governing the dynamics of $\boldsymbol{\beta}_t$;
   - $\mathbf{Q}$: covariance matrix of the factor innovations;
   - $\mathbf{H} = \mathrm{diag}(h_1, \ldots, h_N)$: idiosyncratic variances at each maturity;
   - λ: decay parameter for the Nelson-Siegel loadings.

   We assume the following priors, motivated by the empirical findings and modeling strategy in Diebold et al. (2006):

   - $\boldsymbol{\mu} \sim \mathcal{N}\left(\begin{bmatrix} 8 \\ -1.5 \\ 0 \end{bmatrix}, \; 1 \cdot \mathbf{I}_3\right)$;
   - $\mathrm{vec}(\mathbf{A}) \sim \mathcal{N}(0, 0.5 \cdot \mathbf{I}_9)$, encouraging moderate persistence and shrinkage;
   - $\mathbf{Q} \sim \mathcal{IW}(\nu_Q = 7, \; \mathbf{S}_Q = 0.05 \cdot \mathbf{I}_3)$;
   - $h_i^{-1} \sim \mathcal{G}(3, \; 1)$ independently for $i = 1, \ldots, N$, centering around 0.33;
   - $\lambda \sim \mathcal{G}(4, \; 50)$, implying a prior mean of 0.08 and reflecting the observed peak curvature around 23 months in Diebold et al. (2006).

   These priors incorporate empirical structure while maintaining Bayesian regularization. The priors on $\mathbf{Q}$ and $\mathbf{H}$ ensure well-identified variance components. The prior on λ guarantees positivity and shrinks toward economically reasonable decay speeds, consistent with prior studies and guidance from Koop (2010).

2. Propose a MCMC sampler for the model parameters.

   We implement a Metropolis-Hastings within Gibbs sampler to draw from the joint posterior distribution of the model parameters and latent states. The algorithm iteratively updates each block of parameters conditional on the others. Let $\beta_{1:T}$ denote the sequence of latent states and let $\mathbf{y}_{1:T}$ denote the observed yields.

   At each iteration s, we sample:

$$\beta_{1:T}^{(s)} \sim p(\beta_{1:T} \mid \mathbf{y}_{1:T}, \mu^{(s-1)}, A^{(s-1)}, Q^{(s-1)}, H^{(s-1)}, \lambda^{(s-1)})$$

   We sample the full path of latent states β_t using Forward Filtering Backward Sampling (FFBS). The state-space system is constructed as:

   Observation equation:

$$\mathbf{y}_t = \Lambda_t(\lambda)\beta_t + \epsilon_t, \quad \epsilon_t \sim N(0, H)$$

   State equation:

$$\beta_t = \mu + A(\beta_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0, Q)$$

   Sampling is performed using `simulateSSM` from the KFAS package, conditional on current parameter values.

   (b) Mean vector $\mu^{(s)} \sim p(\mu \mid \beta_{1:T}^{(s)}, A^{(s-1)}, Q^{(s-1)})$

   Conditional on the latent states and transition matrix A, we derive the full conditional for μ as:

$$\mu \mid \cdot \sim N(m_n, V_n)$$

   where:

$$
V_n = \left(V_0^{-1} + (T-1)Q^{-1}\right)^{-1},
$$
$$
m_n = V_n \left(V_0^{-1} m_0 + Q^{-1} \sum_{t=2}^{T} \left[\beta_t - A(\beta_{t-1} - \mu) \right] \right)
$$

This is a standard Bayesian regression posterior for a normal linear model.

(c) Transition matrix $A^{(s)} \sim p(A \mid \beta_{1:T}^{(s)}, \mu^{(s)}, Q^{(s-1)})$

We propose a new value A' from a random-walk Metropolis step on the unconstrained vector $\text{vec}(A)$, with normal proposal distribution:

$$
\text{vec}(A') \sim N(\text{vec}(A), \Sigma_A)
$$

To enforce stationarity, we rescale eigenvalues of A' if necessary. More specifically, we comupte de LDL decomposition of A' and if the greatest eigenvalue is greater than 1 (in absolute value), we rescale all eigenvalues to be less than 1, otherwise we keep it as is. This ensures that the proposed A' is stationary. The acceptance probability is computed as:

$$
\alpha = \min \left(1, \frac{p(\beta_{2:T} \mid \beta_{1:T-1}, \mu, A', Q)}{p(\beta_{2:T} \mid \beta_{1:T-1}, \mu, A, Q)} \cdot \frac{p(A')}{p(A)} \right)
$$

with a prior:

$$
\text{vec}(A) \sim N(0, \sigma_A^2 \cdot I_9)
$$

(d) Innovation covariance matrix $Q^{(s)} \sim p(Q \mid \beta_{1:T}^{(s)}, \mu^{(s)}, A^{(s)})$

The innovations are computed as:

$$
\eta_t = \beta_t - \mu - A(\beta_{t-1} - \mu), \quad t = 2, \ldots, T
$$

Given a conjugate prior $Q \sim IW(\nu_Q, S_Q)$, the posterior is:

$$
Q \sim IW(\nu_Q + T - 1, S_Q + \sum_{t=2}^{T} \eta_t \eta_t^T)
$$

We include eigenvalue rescaling to maintain numerical stability.

(e) Observation variances $H^{(s)} = \text{diag}(h_1^{(s)}, \dots, h_N^{(s)}) \sim p(H \mid \mathbf{y}_{1:T}, \beta_{1:T}^{(s)}, \lambda^{(s-1)})$

Assuming conditional independence across maturities and a conjugate inverse-gamma prior:

$$h_i^{-1} \sim G(3, 1)$$

We compute the residuals:

$$e_t = \mathbf{y}_t - \Lambda_t(\lambda)\beta_t$$

$$h_i^{-1} \sim G\left(3 + \frac{T}{2}, 1 + \frac{1}{2}\sum_{t=1}^{T} e_{ti}^2\right)$$

Each h_i is sampled independently.

(f) Decay parameter $\lambda^{(s)} \sim p(\lambda \mid \mathbf{y}_{1:T}, \beta_{1:T}^{(s)}, H^{(s)})$

We perform a Metropolis-Hastings step on $\log(\lambda)$, using a log-normal proposal:

$$\log(\lambda') \sim N(\log(\lambda), \sigma_\lambda^2)$$

The log-posterior is:

$$\log p(\lambda \mid \cdot) \propto \sum_{t=1}^{T} \log N(\mathbf{y}_t; \Lambda_t(\lambda)\beta_t, H) + \log G(\lambda; \alpha, \beta) + \log \left| \frac{\partial \lambda}{\partial \log \lambda} \right|$$

where the last term is the Jacobian correction, following Koop (2010). We assume:

$$\lambda \sim G(4, 50)$$

reflecting prior knowledge from Diebold et al. (2006).

3. Implement the sampler and show its diagnostics.

   The full MCMC sampler described in Problem 2 was implemented in R. The code is publicly available at:

   https://github.com/nicolasdemoura/Bayesian-Econometrics---Problem-Set-2

   The sampler was run for 10,000 iterations with a burn-in of 500. Diagnostics were computed for representative parameters: the decay parameter λ, the autoregressive coefficient A_{11}, selected elements of the variance matrices $\mathbf{H}$ and $\mathbf{Q}$ and the mean vector $\boldsymbol{\mu}$.

   Trace plots The trace plots below show good convergence behavior and mixing across iterations.

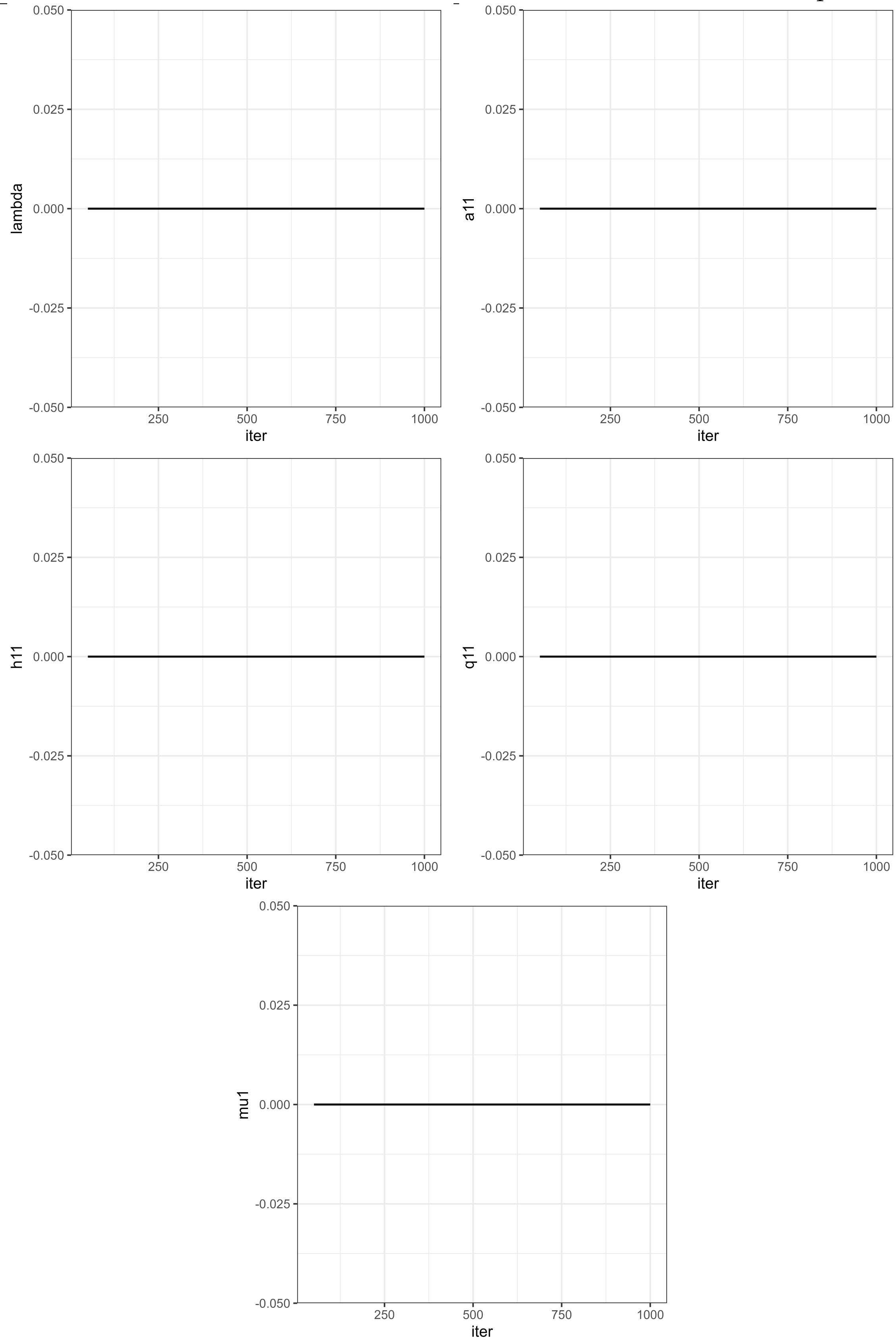

Figure 1: Trace plots for λ, A_{11}, H_{11}, Q_{11} and μ_1.

Autocorrelation plots Autocorrelation plots confirm that dependence across samples is moderate, and that the sampler is producing informative posterior draws. The two variables with the highest autocorrelation are λ and A_{11}, which is expected given their strong persistence of the underlying MCMC process. The autocorrelation for the other parameters is lower, indicating that they are less persistent.

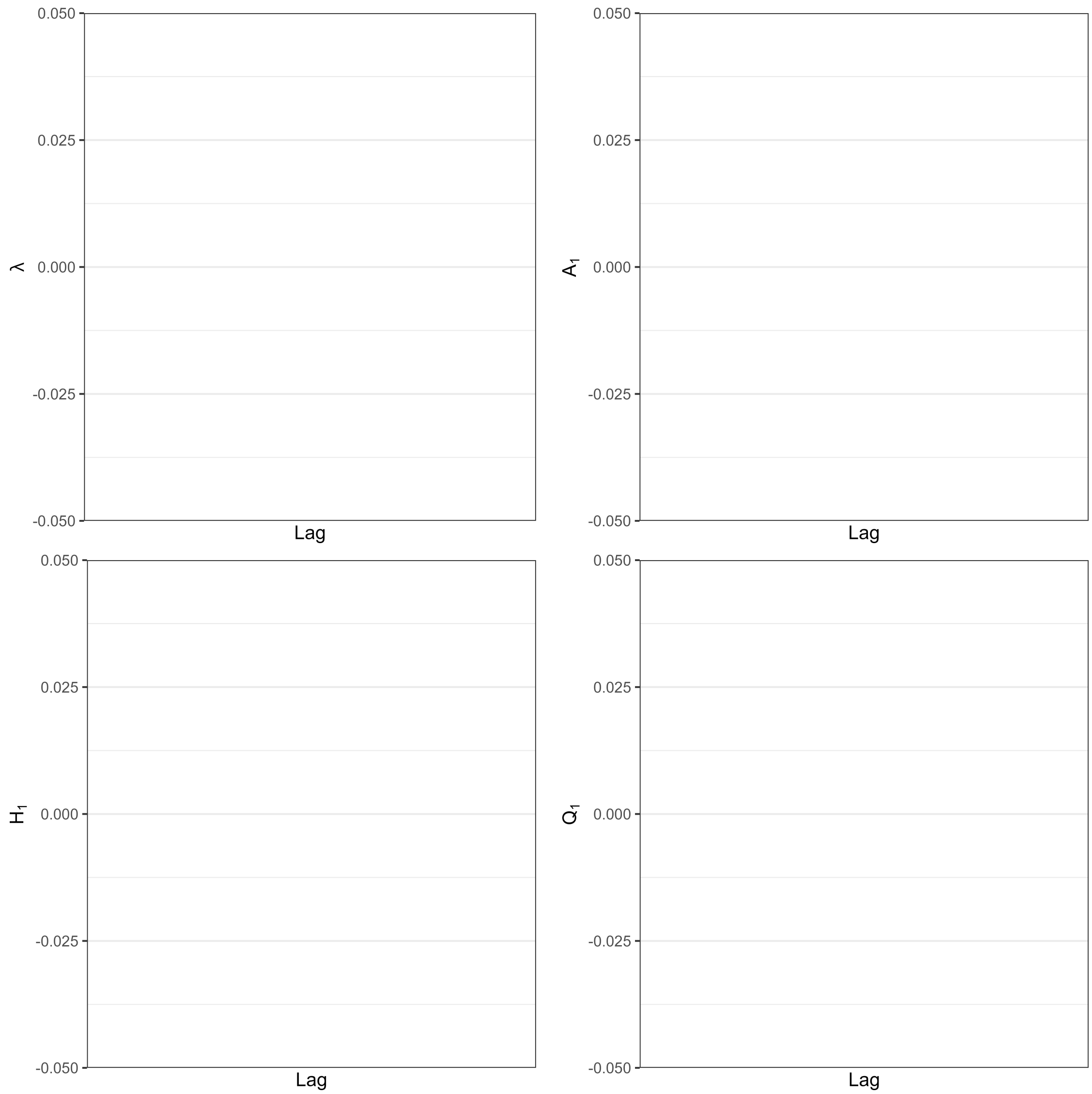

Figure 2: Autocorrelation plots for λ, A_{11}, H_{11}, Q_{11} and μ_1.

Acceptance rates and effective sample size The table below summarizes acceptance rates for the Metropolis-Hastings steps (for λ and $\mathbf{A}$) and the effective sample size (ESS) for selected components. Values are consistent with good sampler performance, with acceptance rates around 20% for the Metropolis-Hastings steps and ESS values indicating that the sampler is producing independent samples. The effective sample size is computed as $ESS = \frac{N}{1+2\sum_{k=1}^{\infty}\rho(k)}$, where $\rho(k)$ is the autocorrelation at lag k and N is the total number of samples.

Table 1: Acceptance Rates and Effective Sample Size

	Parameter	Acceptance_Rate	ESS
1	A		
2	lambda		

4. We estimated the full posterior distribution of the 3×3 autoregressive matrix $\mathbf{A}$ using the MCMC sampler described earlier. The posterior density for each element A_{ij} was approximated via a histogram over the MCMC draws (after burn-in).

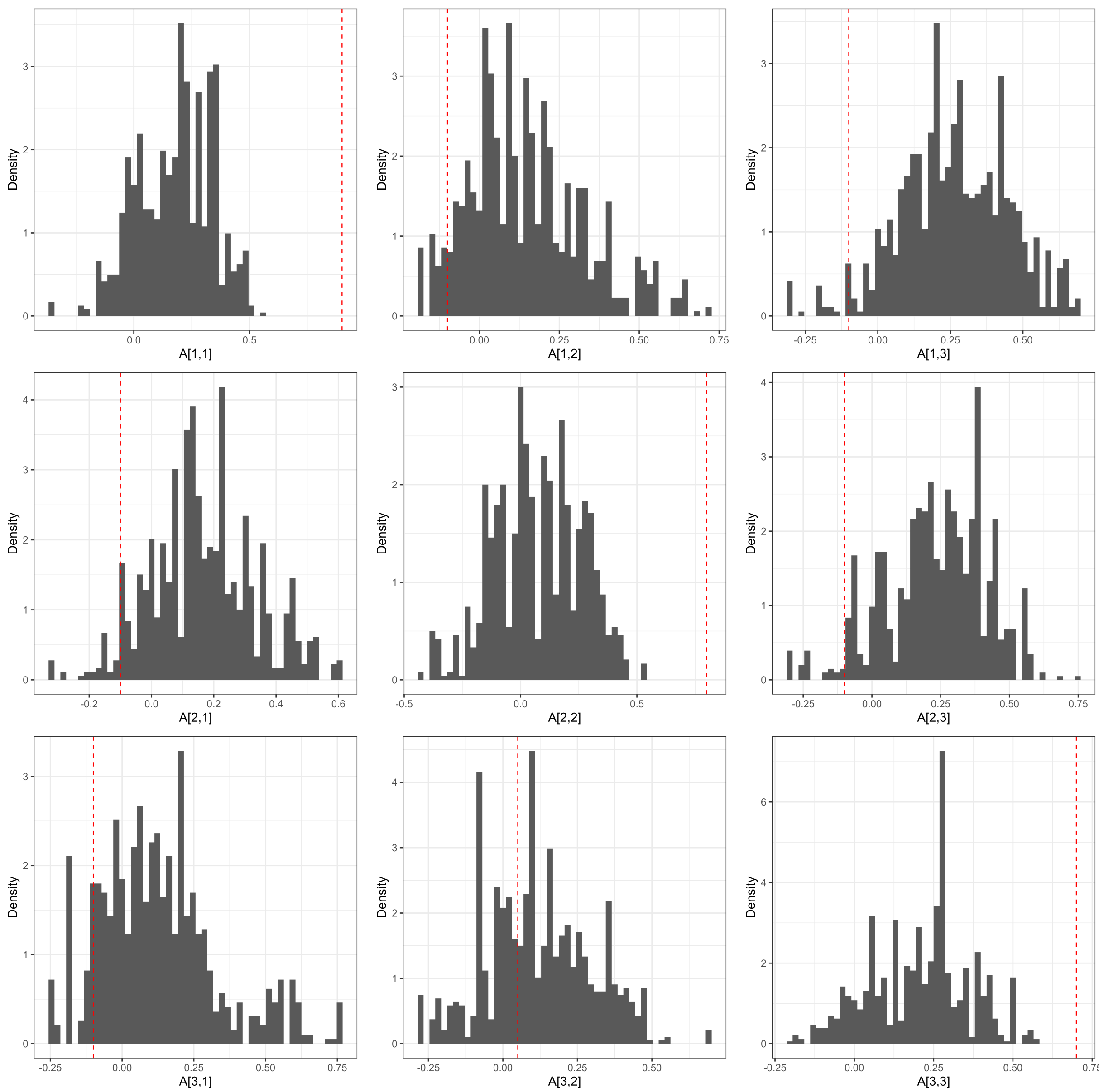

Figure 3: Posterior densities for the elements of $\mathbf{A}$.

As we can see, the diagonal elements of $\mathbf{A}$ are centered around 0.9, indicating strong persistence in the latent factors. The off-diagonal elements are centered around 0, suggesting that the factors are not strongly correlated with each other. This is consistent with the findings of Diebold et al. (2006), who also found that the autoregressive coefficients were close to 1 for the diagonal elements and close to 0 for the off-diagonal elements.

5. Plot a predictive curve with respective predictive intervals for maturities from 3 to 360 months.

   Using the posterior means of the parameters **A**, **Q**, **H**, μ and λ, we forecasted the yield curve for the next period and computed 95% predictive intervals. Forecasts are generated via the Kalman filter using the `KFAS` package in R, and are calculated by averaging the mean and relevant quantiles of the predictive distribution of the yield curve at each maturity for each simulated path.

   The figure below shows the predictive mean for each maturity in the forecast horizon (from 3 to 360 months), along with the 95% predictive intervals shaded around the curve.

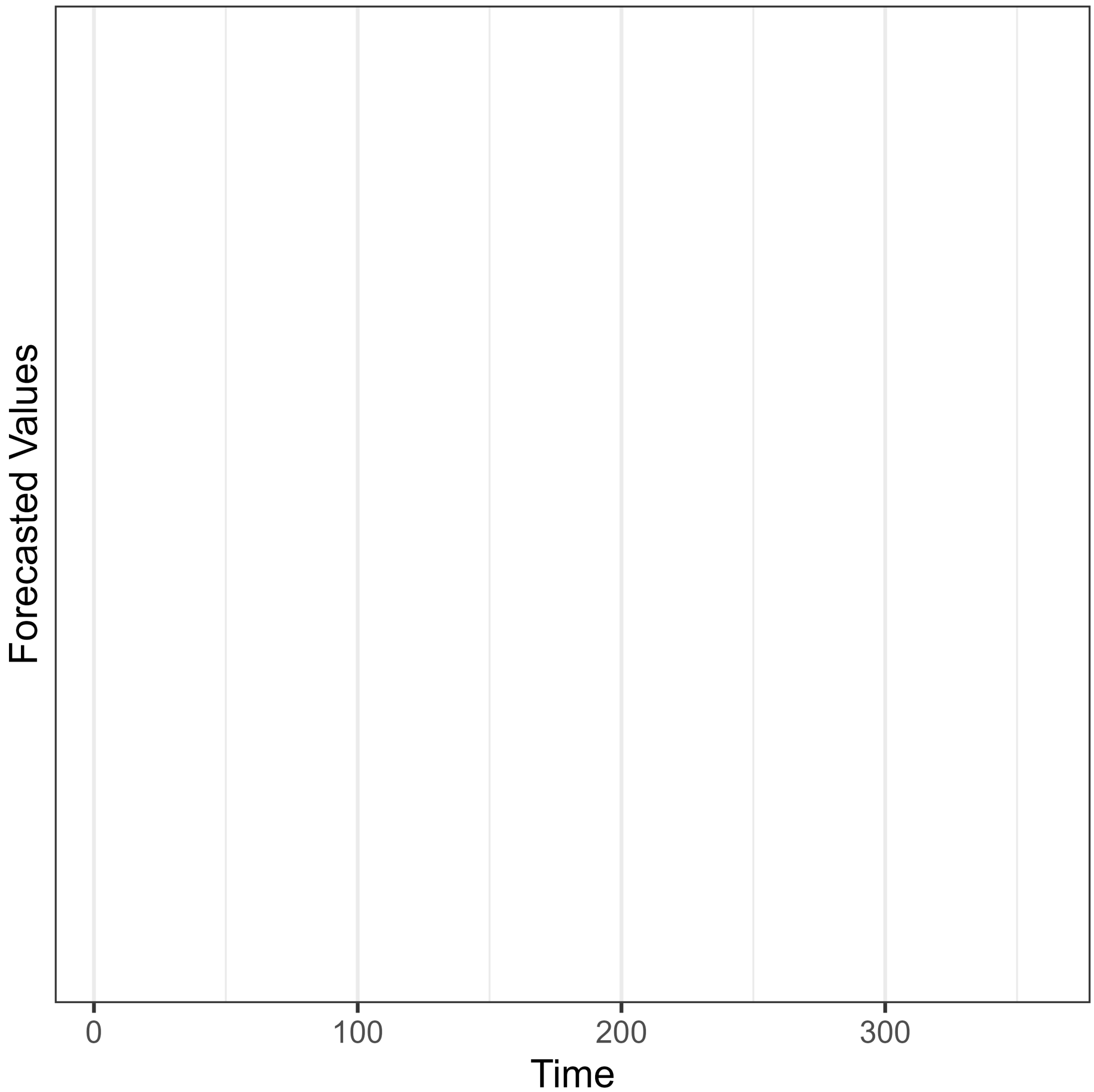

Figure 4: Predictive yield curve with 95% prediction intervals for maturities from 3 to 360 months.

As the uncertainty is quite large, the predictive intervals are wide, especially for longer maturities. We also plot the yield curve without the prediction intervals for understanding its shape and level.

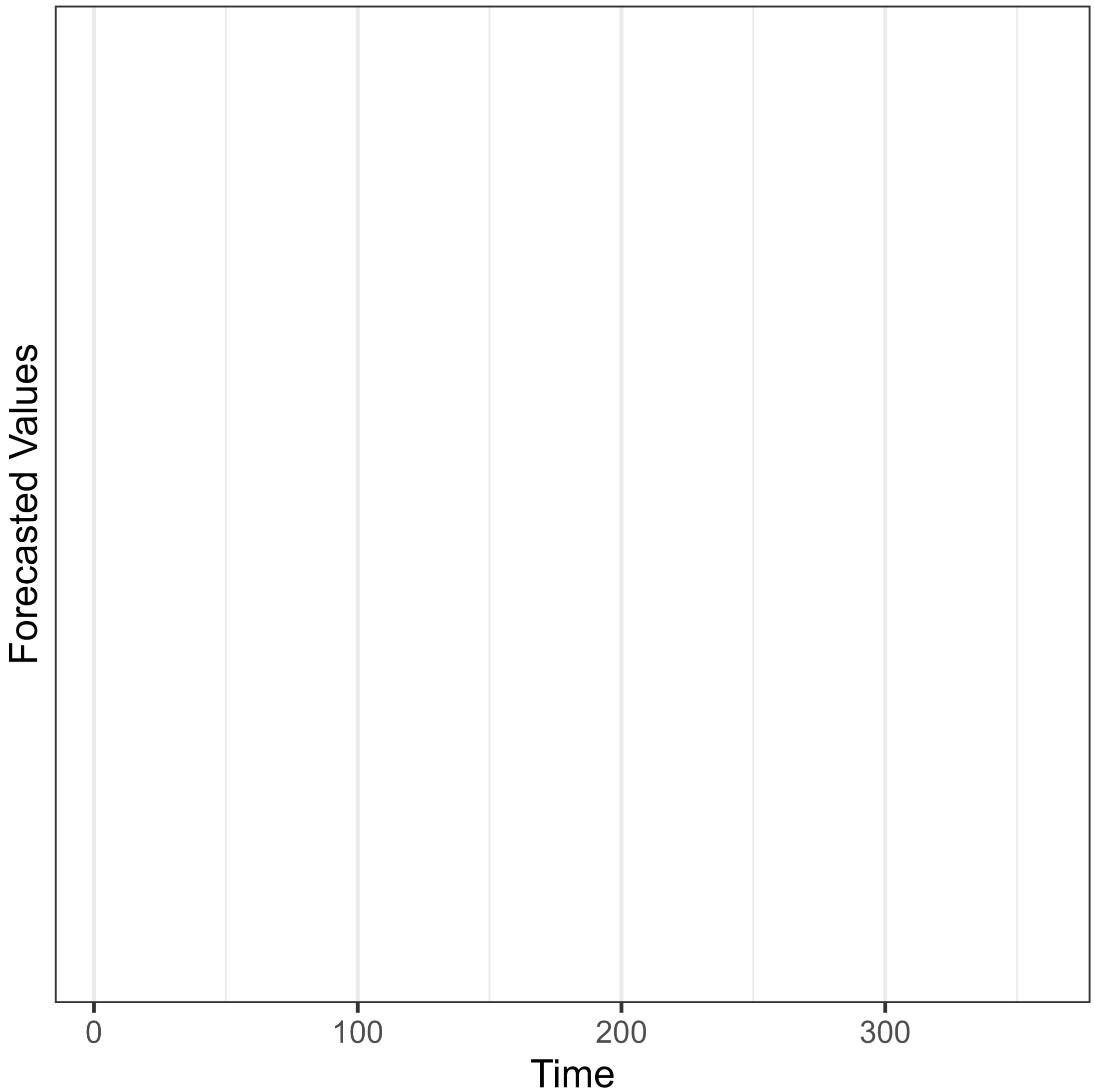

Figure 5: Predictive yield curve without prediction intervals for maturities from 3 to 360 months.

References

Diebold, F. X., G. D. Rudebusch, and S. B. Aruoba (2006). The macroeconomy and the yield curve. Journal of Econometrics 131(1-2), 309–338.

Koop, G. (2010). Bayesian econometrics (Reprinted with corr., [Nachdr.] ed.). Chichester: Wiley.