

PROBLEM SET # 2

The term structure of interest rates, known as the yield curve, illustrates the connection between the remaining time until debt securities mature and the yield they offer. Yield curves serve multiple practical purposes, such as pricing different fixed-income securities. They are closely monitored by both market participants and policymakers, as they can provide valuable insights into the market's sentiment regarding the future direction of the policy rate and the overall macroeconomic outlook.

The Nelson-Siegel model, introduced in ?, interpolates the yield curve (in terms of spot rates) by the following function:

$$s(\tau) = \beta_0 + \beta_1 \left\{ \frac{1 - e^{-\tau/\lambda}}{\tau/\lambda} \right\} + \beta_2 \left[\left\{ \frac{1 - e^{-\tau/\lambda}}{\tau/\lambda} \right\} - e^{-\tau/\lambda} \right], \quad (1)$$

where $s(\tau)$ is the spot rate at any given time to maturity τ , and $\beta_0, \beta_1, \beta_2$, and λ are model parameters. This parametric model captures stylized shapes observed in yield curves, including monotonic, humped, and S-shaped curves. Parameters β_0, β_1 , and β_2 are interpreted as the level, slope, and curvature factors, respectively, as they control the long, short, and medium segments of the curve. The decay parameter λ determines the exponential decay rate (in months to maturity) of the slope and curvature factors, in addition to controlling the location of the hump (or trough) associated with the curvature factor.

We construct the term structure of interest rates by taking weekly closing prices of zero-coupon yields released by Bloomberg. We consider the following fixed maturities (τ): 3, 6, 12, 24, 30, 36, 48, 60, 72, 84, 96, 108, 180, 240, and 360 months. The data covers three years, from the first week of 2019 until the last week of 2021.

Let y_{it} denote the weekly yield on date t ($t = 1, \dots, 157$) at maturity τ_i ($i = 1, \dots, 15$), i.e., $t = 1$ is the 1st week of 2019, $t = 157$ is the last week of 2021, and $\tau_1 = 3, \tau_2 = 6, \dots, \tau_{15} = 360$. The yield is modeled by

$$y_{it} = s(\tau_i) + u_{it}, \quad u_{it} \sim \text{NID}(0, \sigma^2). \quad (2)$$

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

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

- $\mu = (\mu_1, \mu_2, \mu_3)$: unconditional means of the factors;
- \mathbf{A} : a 3×3 autoregressive matrix governing the dynamics of β_t ;
- \mathbf{Q} : covariance matrix of the factor innovations;
- $\mathbf{H} = \text{diag}(h_1, \dots, h_N)$: 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 ?:

- $\mu \sim \mathcal{N}\left(\begin{bmatrix} 8 \\ -1.5 \\ 0 \end{bmatrix}, 100 \cdot \mathbf{I}_3\right)$;
- $\text{vec}(\mathbf{A}) \sim \mathcal{N}(0, 0.5 \cdot \mathbf{I}_9)$, encouraging moderate persistence and shrinkage;
- $\mathbf{Q} \sim \mathcal{IW}(v_Q = 7, \mathbf{S}_Q = 0.05 \cdot \mathbf{I}_3)$;
- $h_i^{-1} \sim \mathcal{G}(3, 1)$ independently for $i = 1, \dots, N$, 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 ?.

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 ?.

2. Propose a MCMC sampler for the model parameters.

※ MCMC Sampling Algorithm

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:

$$V_n = (V_0^{-1} + (T-1)Q^{-1})^{-1},$$

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

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. 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:

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

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. We assume:

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

reflecting prior knowledge from ?.

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} , and selected elements of the variance matrices \mathbf{H} and \mathbf{Q} .

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

Autocorrelation plots Autocorrelation plots confirm that dependence across samples is moderate, and that the sampler is producing informative posterior draws.

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.

Table 1: Acceptance Rates and Effective Sample Size

	Parameter	Acceptance_Rate	ESS
1	A	0.245	6.426
2	lambda	0.940	0.899

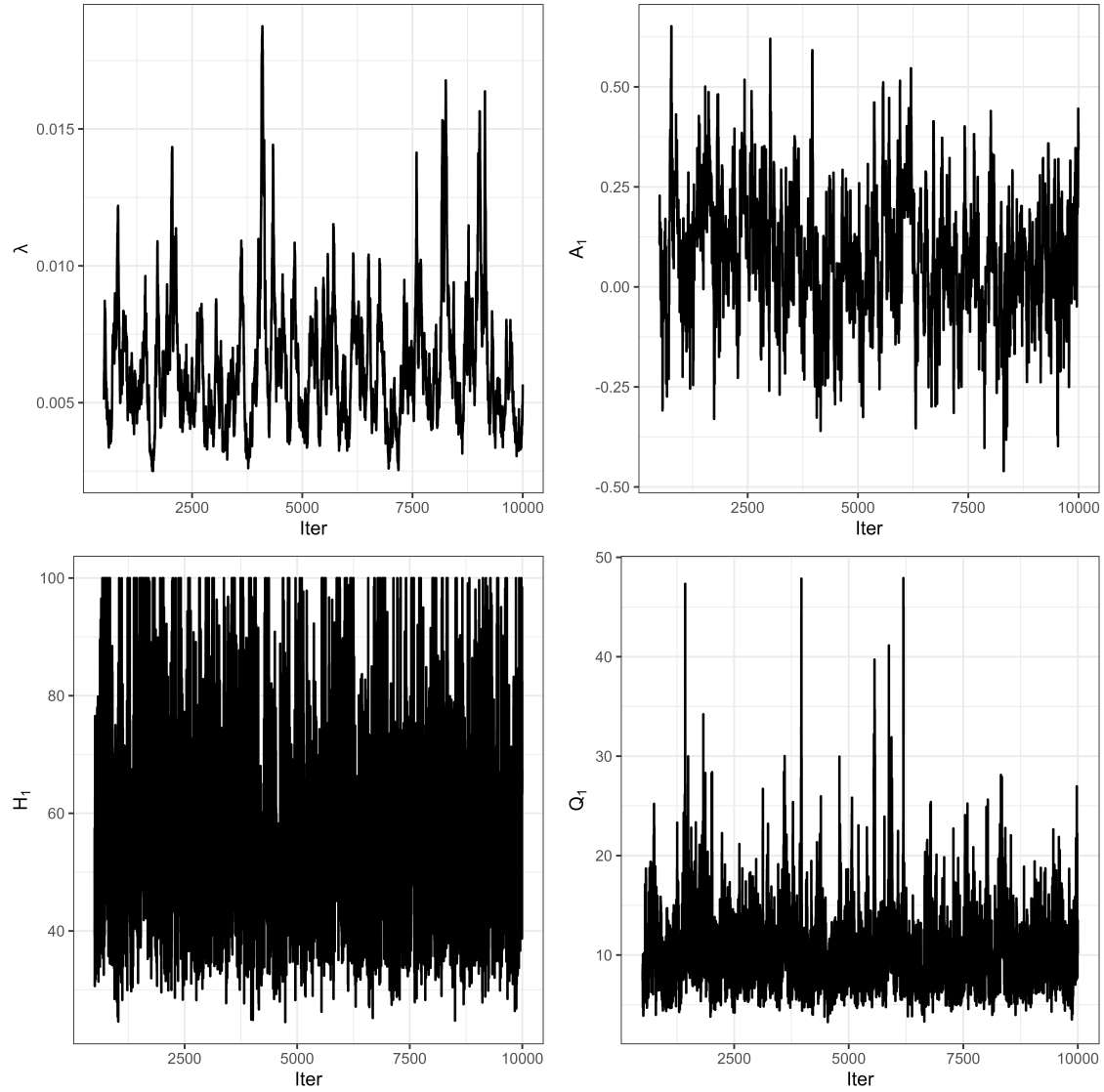


Figure 1: Trace plots for λ , A_{11} , H_{11} and Q_{11} .

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).

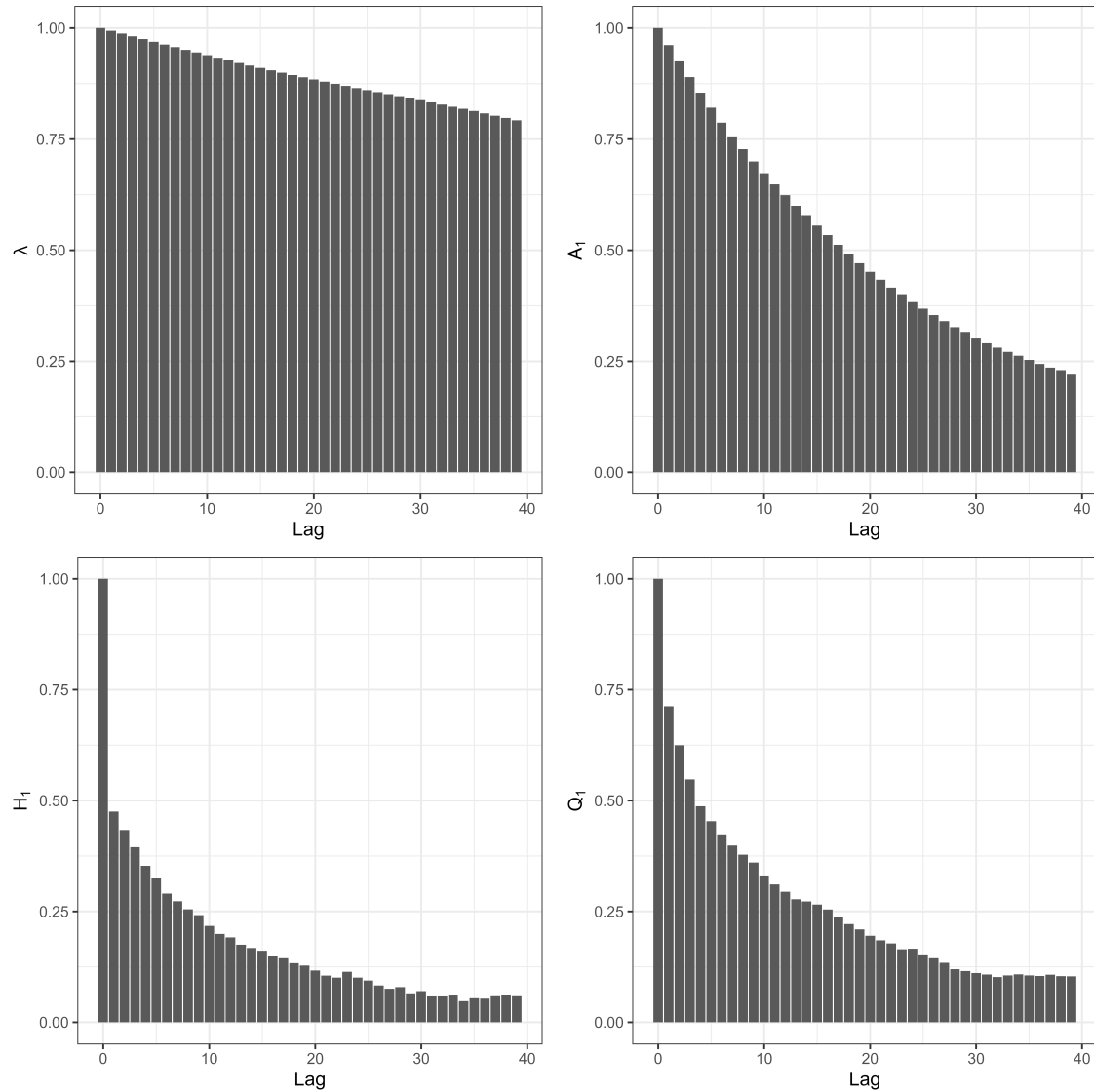


Figure 2: Autocorrelation plots for λ , A_{11} , H_{11} and Q_{11} .

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

Using the posterior means of the parameters \mathbf{A} , \mathbf{Q} , \mathbf{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

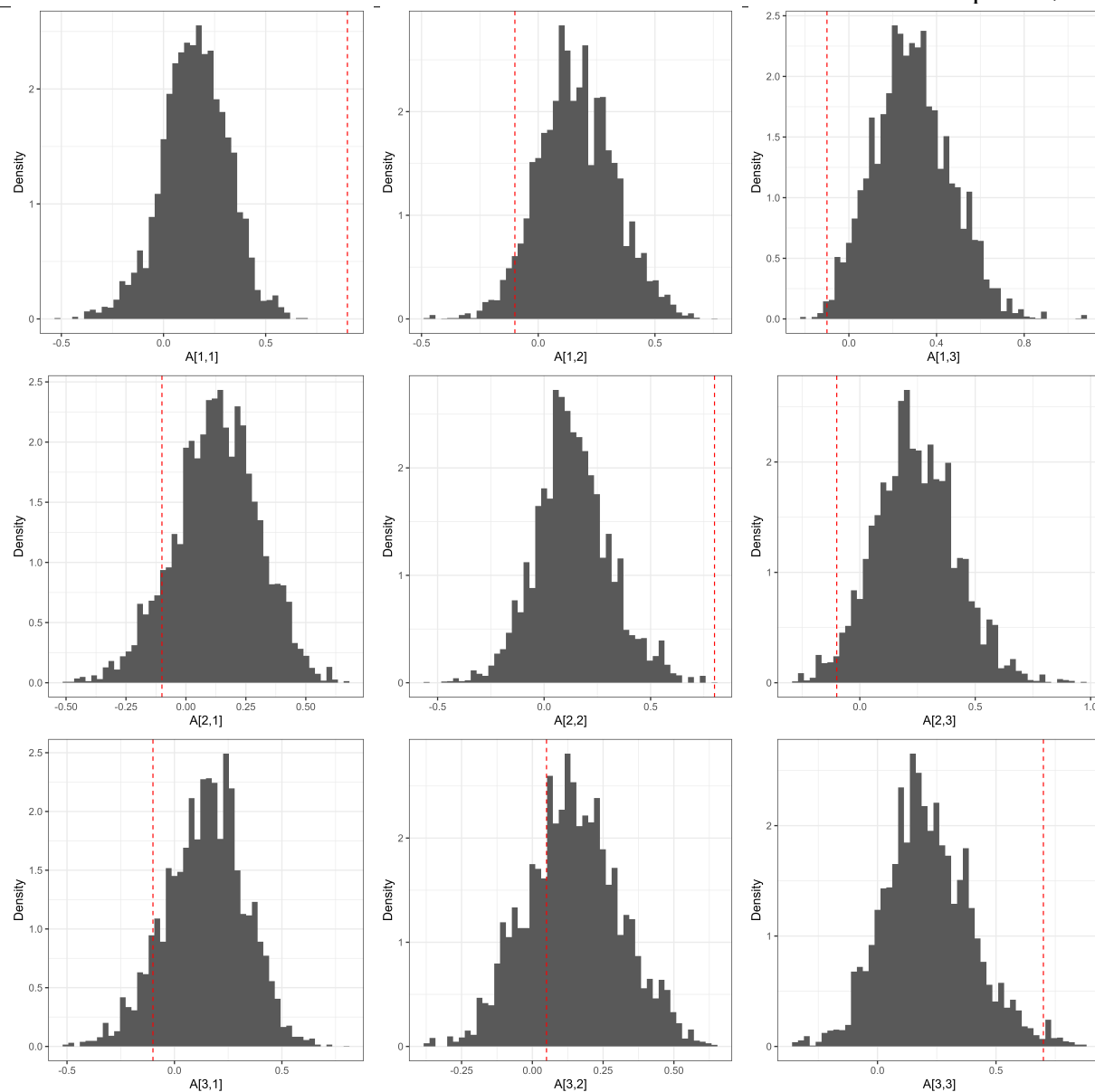


Figure 3: Posterior densities for the elements of \mathbf{A} .

curve.

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.

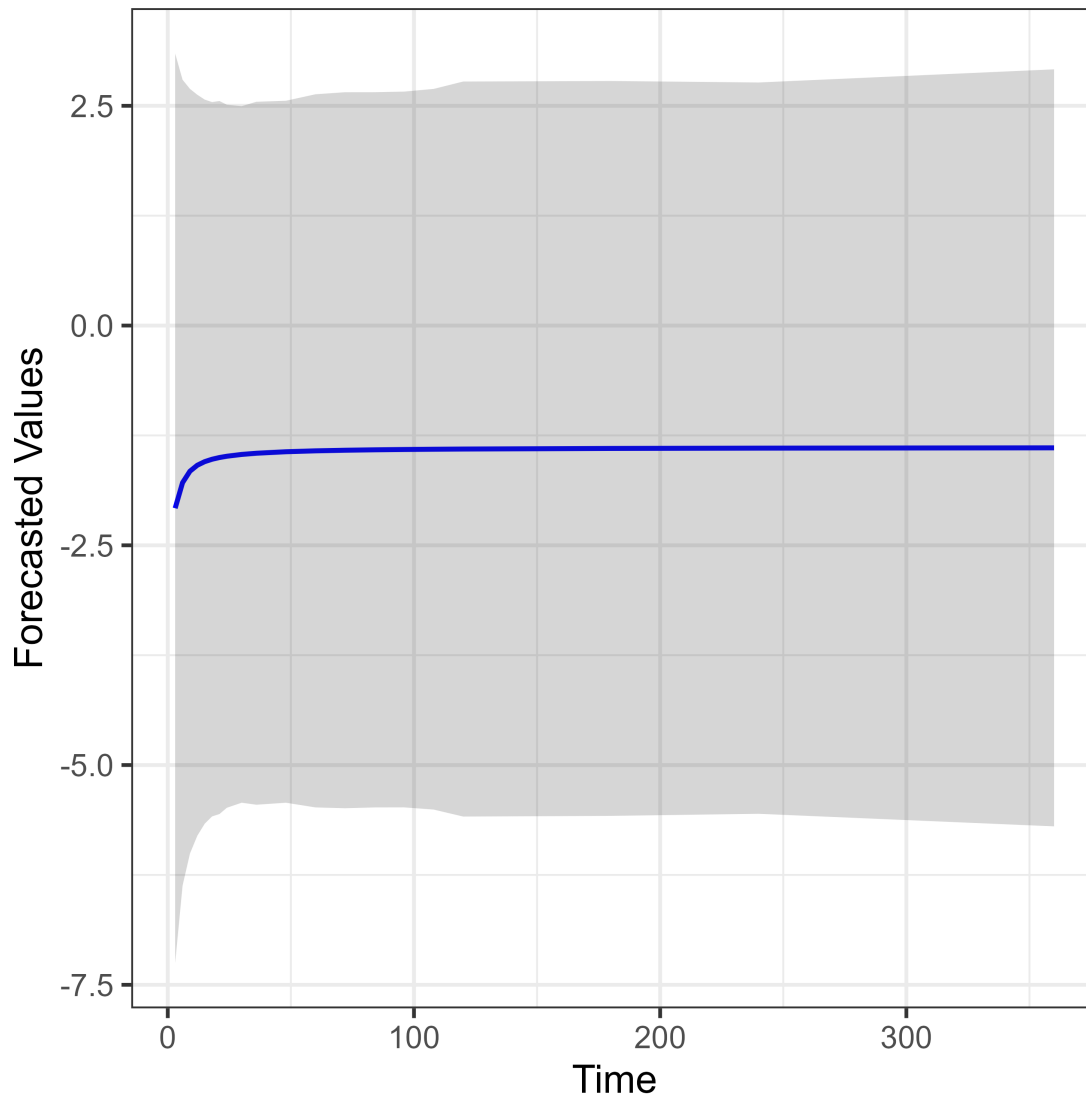
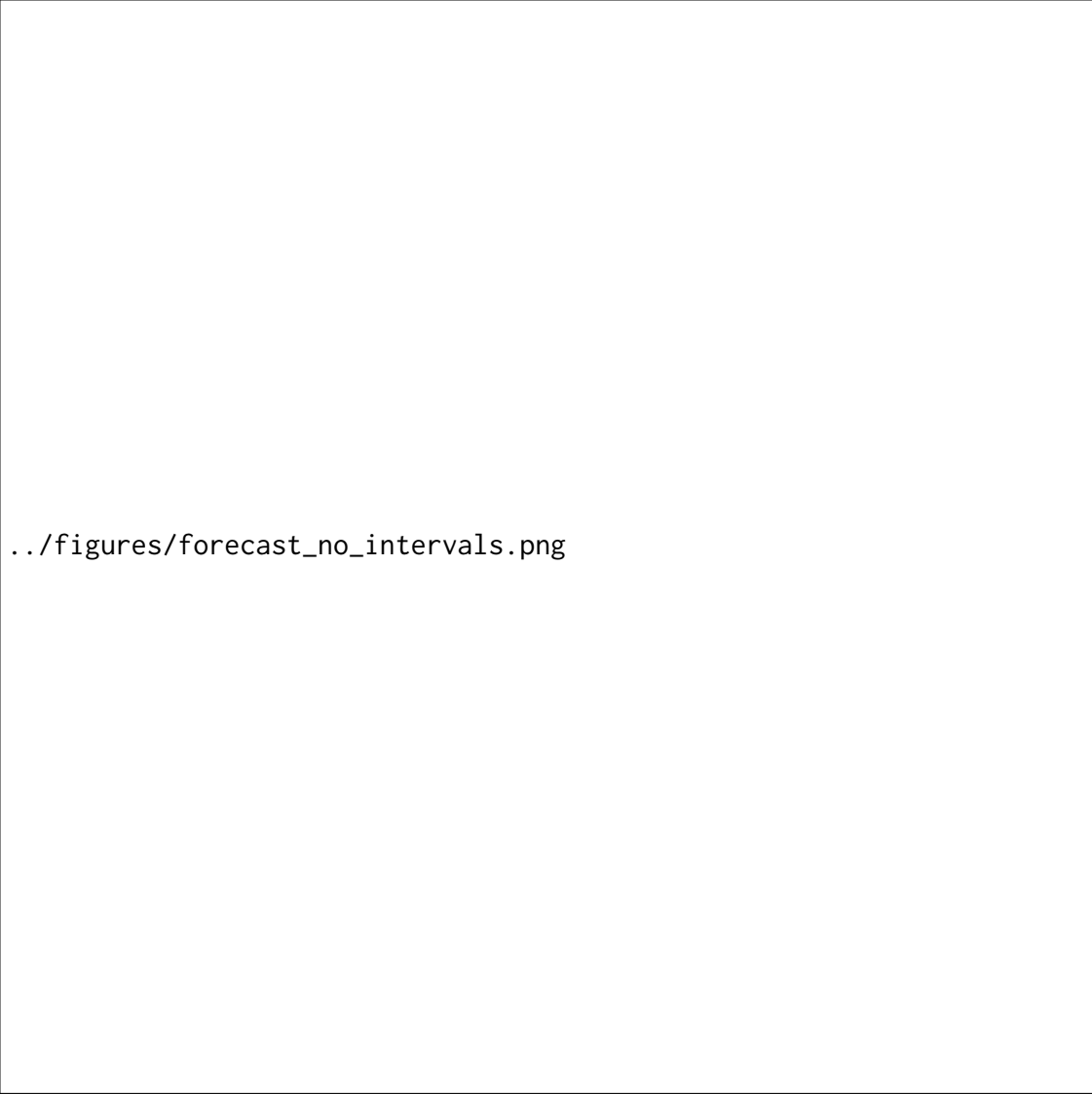


Figure 4: Predictive yield curve with 95% prediction intervals for maturities from 3 to 360 months.

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*. John Wiley & Sons.
- Nelson, C. R. and A. F. Siegel (1987, January). Parsimonious Modeling of Yield Curves. *The Journal of Business* 60(4), 473.



../figures/forecast_no_intervals.png

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