Bayesian Inference for a Stochastic Volatility Nelson-Siegel Model

Nikolaus Hautsch

Humboldt-Universität zu Berlin, Institute for Statistics and Econometrics

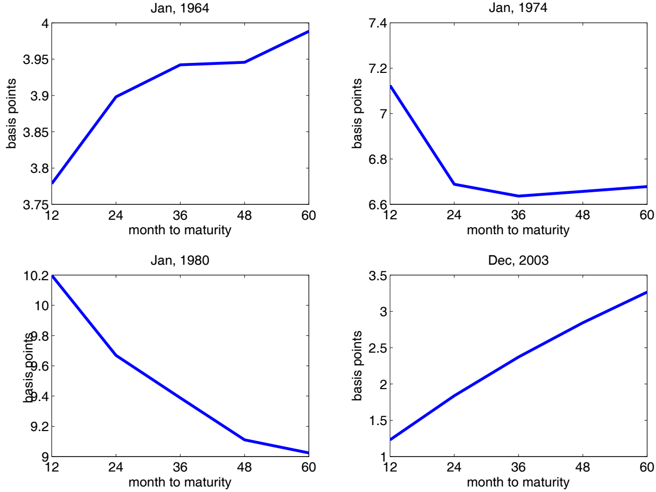
Fuyu Yang

Humboldt-Universität zu Berlin, Institute for Statistics and Econometrics

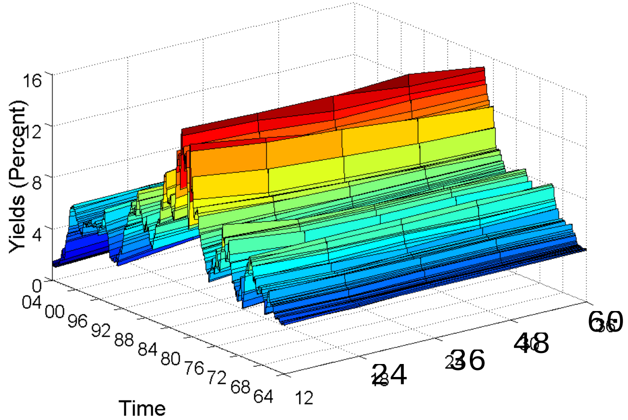
Feb 12, 2010

mailto:fuyu.yang@cms.hu-berlin.de

Different Shapes of Yield Curves



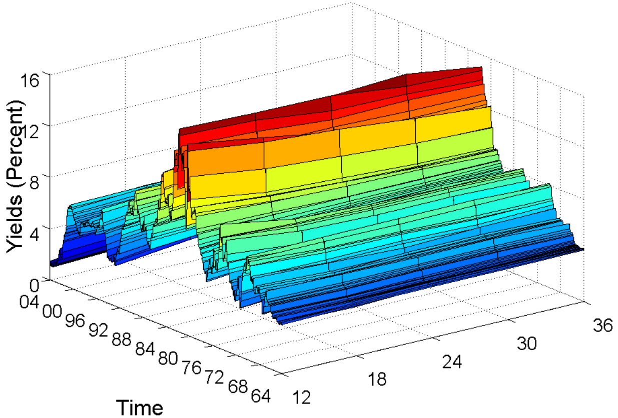
Yield Data Plot



Can we forecast yields?

- Use a parsimonious Nelson Siegel model to forecast yields

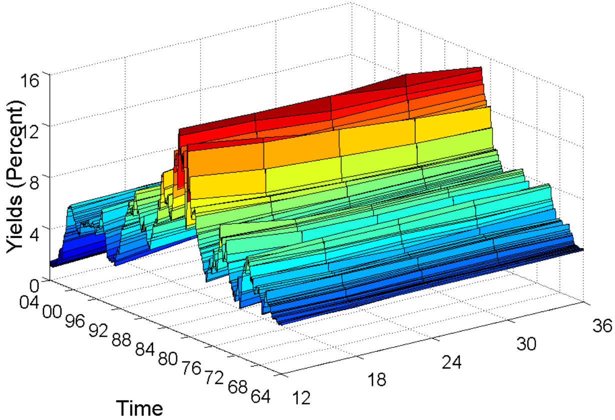
Yield Data Plot



How can we capture the time varying variances in all yields?

- We propose a Stochastic Volatility Nelson Siegel model.

Yield Data Plot



How we do density forecast for yields?

- Can be achieved in a Bayesian framework without extra cost.

Why we care about yield curve forecasts?

- For monetary policy makers:
  - The yield curve carries information about prospective evolution of economic activity, inflation etc. e.g. an inverted yield curve indicates that investors expect the economy to slow down in the future.
  - Interest rate spreads (difference between 10 year and 3 month rates, or 10 year and 2 year rates) are predictive indicators of recession and inflation e.g. steep yield curve can be observed at the beginning of economic expansion (after recession).
- For investors...

Outline

1. Model Specifications
2. Estimation in Bayesian Framework
3. Forecasting in Bayesian Framework
4. Empirical Results
5. Conclusion

Outline

Nelson Siegel Model (1987)

$$y_t\left(n\right) = f_{1,t} + f_{2,t}\left[\frac{1-e^{-\lambda\cdot n}}{\lambda\cdot n}\right] + f_{3,t}\left[\frac{1-e^{-\lambda\cdot n}}{\lambda\cdot n} - e^{-\lambda\cdot n}\right]$$

- $y_t\left(n\right)$: yields at time t, with n period to maturity.

Diebold Li Nelson Siegel Model (2006)

- $[f_{1,t}, f_{2,t}, f_{3,t}]$: level, slope, curvature factors
- $\left[\ 1,\quad \frac{1-e^{-\lambda\cdot n}}{\lambda\cdot n},\quad \frac{1-e^{-\lambda\cdot n}}{\lambda\cdot n} - e^{-\lambda\cdot n}\ \right]$: factor loadings.

Denote: $\mathbf{y}_t = \{y_t\left(1\right), y_t\left(2\right), \ldots, y_t\left(N\right)\}'$, $\mathbf{f}_t = \{f_{1,t}, f_{2,t}, f_{3,t}\}'$.

$$\mathbf{y}_t = \mathbf{A}\mathbf{f}_t + \varepsilon_t \tag{1}$$

$$\mathbf{f}_t - \mu_{\mathbf{f}} = \Phi_{\mathbf{f}}\left(\mathbf{f}_{t-1} - \mu_{\mathbf{f}}\right) + \mathbf{u}_t \tag{2}$$

where $\varepsilon_t \overset{i.i.d}{\sim} f_{MN}\left(0_N, H_\varepsilon\right)$, and $\mathbf{u}_t \overset{i.i.d}{\sim} f_{MN}\left(0_N, H_u\right)$ and

$$\mathbf{A} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\cdot 1}}{\lambda\cdot 1} & \frac{1-e^{-\lambda\cdot 1}}{\lambda\cdot 1} - e^{-\lambda\cdot 1} \\ 1 & \frac{1-e^{-\lambda\cdot 2}}{\lambda\cdot 2} & \frac{1-e^{-\lambda\cdot 2}}{\lambda\cdot 2} - e^{-\lambda\cdot 2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\cdot n}}{\lambda\cdot n} & \frac{1-e^{-\lambda\cdot n}}{\lambda\cdot n} - e^{-\lambda\cdot n} \end{pmatrix} : N \times 3$$

Stochastic Volatility Nelson Siegel Model (2010)

$$\mathbf{f}_t - \mu_{\mathbf{f}} = \Phi_{\mathbf{f}} \left(\mathbf{f}_{t-1} - \mu_{\mathbf{f}} \right) + \mathbf{u}_t \tag{3}$$

$$\mathbf{u}_t = \begin{pmatrix} \sqrt{\exp h_{1,t}} & 0 & 0 \\ 0 & \sqrt{\exp h_{2,t}} & 0 \\ 0 & & \sqrt{\exp h_{3,t}} \end{pmatrix} \times \zeta_t \tag{4}$$

$$f_{j,t} - \mu_{j,f} = \phi_{j,f} \left(f_{j,t-1} - \mu_{j,f} \right) + \exp \left(\frac{h_{j,t}}{2} \right) \zeta_{j,t} \tag{5}$$

$$h_{j,t} - \mu_{j,h} = \phi_{j,h} \left(h_{j,t-1} - \mu_{j,h} \right) + \sigma_j \epsilon_{j,t} \tag{6}$$

where $j = 1, 2, 3$, $\zeta_{j,t} \overset{i.i.d}{\sim} f_N(0,1)$ and $\epsilon_{j,t} \overset{i.i.d}{\sim} f_N(0,1)$

Outline

1. Model Specifications

2. Estimation in Bayesian Framework

3.Forecasting in Bayesian Framework

4. Empirical Results

5.Conclusion

Estimation

Estimating the model is challenging

- Parameters have to be estimated:
  $\theta = \{\lambda, H_\varepsilon, [\mu_{j,f}, \phi_{j,f}], [\mu_{j,h}, \phi_{j,h}, \sigma_j]\}$ and the latent variables f_j and h_j for $j = 1, 2, 3$.
- The latent yield factors f and log Volatilities h are highly correlated, blocking f and h is needed (Chib and Greenberg 1996)
- An efficient MCMC algorithm is a solution.

Blocking the highly correlated latent variables

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{A}\mathbf{f}_t + \varepsilon_t \\
\mathbf{f}_t - \mu_{\mathbf{f}} &= \Phi_{\mathbf{f}}\left(\mathbf{f}_{t-1} - \mu_{\mathbf{f}}\right) + \mathbf{u}_t \\
f^*_{j,t} &= h_{j,t} + z_{j,t} \\
h_{j,t+1} - \mu_{j,h} &= \phi_{j,h}\left(h_{j,t} - \mu_{j,h}\right) + \sigma_j \epsilon_{j,t} : j = 1, 2, 3
\end{aligned}
$$

where

$$
\begin{aligned}
f^*_{j,t} &= \ln\left\{\left[f_{j,t} - \mu_{j,f} - \phi_{j,f}\left(f_{j,t-1} - \mu_{j,f}\right)\right]^2 + c\right\} \\
z_{j,t} &= \ln\left(\zeta^2_{j,t}\right)
\end{aligned}
$$

Remarks

1. c is the Fuller offset parameter, we choose $c = 0.0001$.
2. The SVNS model can be fitted in the state space form.
3. Approximate the distribution of $z_{j,t}$ by 7 mixture of normal densities (Kim, Chib,Shephard,1998).

$$z_t \mid s_t \sim f_N\left(m_{s_t}, \nu_{s_t}^2\right)$$

$$\Pr\left(s_t = i\right) = q_i, i \leq 7, t \leq T$$

4. This approximation allows us to draw the whole block $\{h_j\}$
5. The approximation error can be removed at the end of posterior sampling by a re-weighting procedure (Chib, 2002)

MCMC Alogrithm

1. Initialize $\theta = \{\lambda, H_\varepsilon, [\mu_{j,f}, \phi_{j,f}], [\mu_{j,h}, \phi_{j,h}, \sigma_j]\}$ and the latent variables f_j and h_j with $f_{j,0} = \mu_{j,f}$ and $h_{j,0} = \mu_{j,h}$ for $j = 1, 2, 3$.

2. Run a Gibbs sampler for steps (a) - (d) using S replications, where the initial S_0 draws are discarded:

   2.1 Sample H_ε from
   $H_\varepsilon | y, \lambda, [\mu_{j,f}, \phi_{j,f}], [\mu_{j,h}, \phi_{j,h}, \sigma_j], f_j, h_j, \quad j = 1, 2, 3.$

   2.2 Sample λ from $\lambda | y, H_\varepsilon, [\mu_{j,f}, \phi_{j,f}], [\mu_{j,h}, \phi_{j,h}, \sigma_j], f_j, h_j$ using a Griddy-Gibbs sampling method.

   2.3 Sample $[\mu_{j,f}, \phi_{j,f}], f_j$ from $[\mu_{j,f}, \phi_{j,f}], f_j | y, \lambda, H_\varepsilon, h_j$ using a simulation smoother.

2.4 Run (i)-(iv) 3 times to estimate $h_j \mid f_j$, with $j = 1, 2, 3$,respectively:

(i) Compute f_j^* and run the loop (ii)-(iv) 2000 times. Discard the results from the initial 500 loops.

(ii) Sample $s|f_j^*, h_j$.

(iii) Sample $h_j|f_j^*, s, [\mu_{j,h}, \phi_{j,h}, \sigma_j]$ using a simulation smoother.

(iv) Sample $[\mu_{j,h}, \phi_{j,h}, \sigma_j] \,|f_j^*, h_j$ using the Metropolis-Hastings algorithm proposed by Chib et al. (2002) choosing a multivariate t-density as the proposal density.

Outline

Forecasting of Yields

- *h*-step ahead Point forecast

$$\widehat{f}_{t+h} = \widehat{\mu}_f + \widehat{\phi}_f \left(\widehat{f}_t - \widehat{\mu}_f \right), \tag{7}$$

$$\widehat{y}_{t+h} = \widehat{A}\widehat{f}_{t+h}, \tag{8}$$

- *h*-step ahead Density forecast

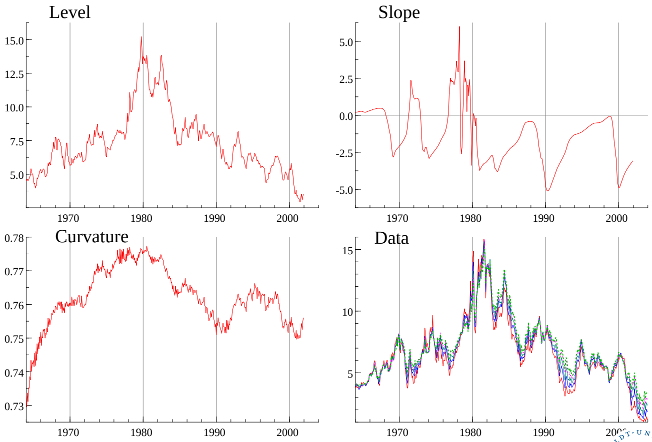
$$\widehat{f}^{(s)}_{t+h} = \widehat{\mu}^{(s)}_f + \widehat{\phi}^{(s)}_f \left(\widehat{f}^{(s)}_t - \widehat{\mu}^{(s)}_f \right) + \widehat{\eta}^{(s)}_{t+h},$$

$$\widehat{y}^{(s)}_{t+h} = \widehat{A}^{(s)} \left(\widehat{f}^{(s)}_{t+h} + \widehat{\eta}^{(s)}_{t+h} \right) + \widehat{\varepsilon}^{(s)}_{t+h},$$

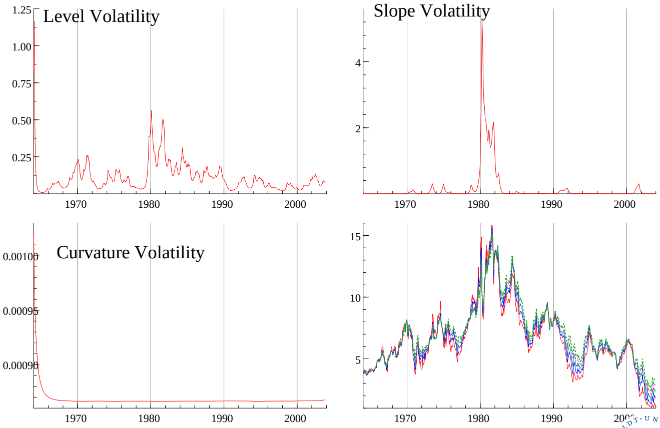
where $\widehat{\eta}^{(s)}_{t+h} = \Sigma^{(s)}_{\eta} u_{\eta,t+h}$ and $\widehat{\varepsilon}^{(s)}_{t+h} = \Sigma^{(s)}_{\varepsilon} u_{\varepsilon,t+h}$, $u_{\varepsilon,t+h}$ and $u_{\eta,t+h}$ are i.i. standard normally distributed.

Outline

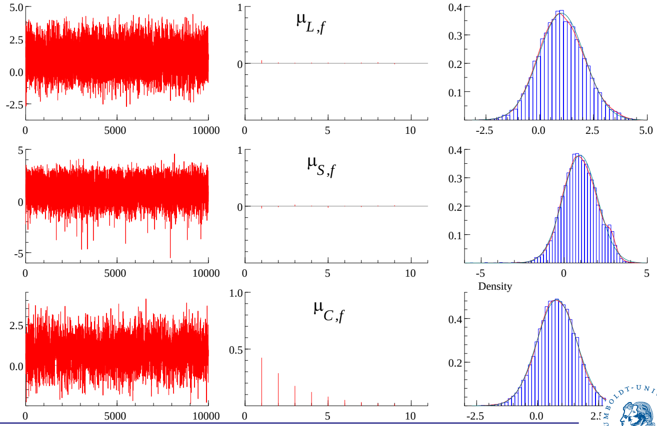
Extracted yield factors f



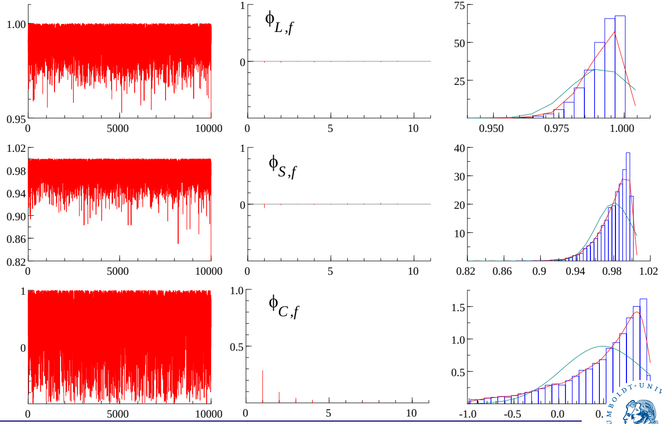
Extracted log Volatilities h



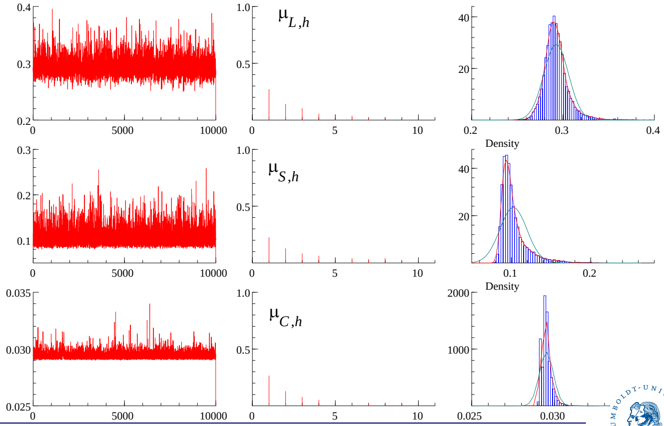
Draws, autocorrelations and posterior densities μ_f



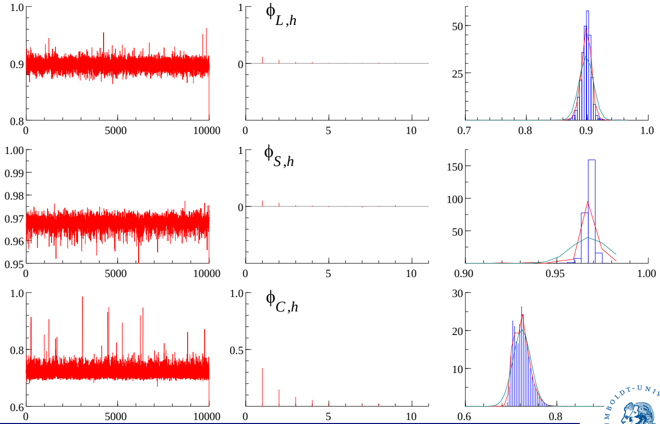
Draws, autocorrelations and posterior densities ϕ_f



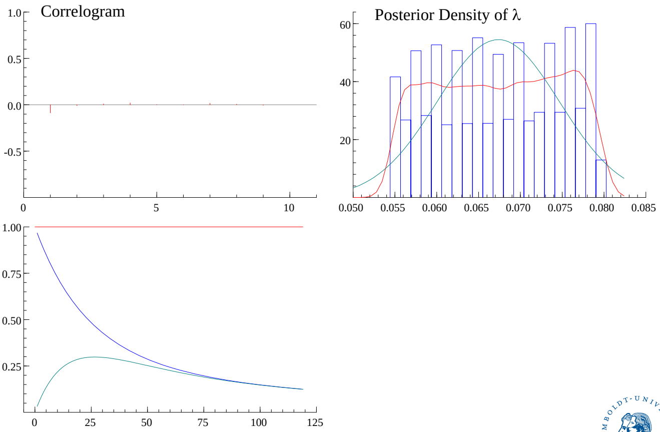
Draws, autocorrelations and posterior densities μ_h



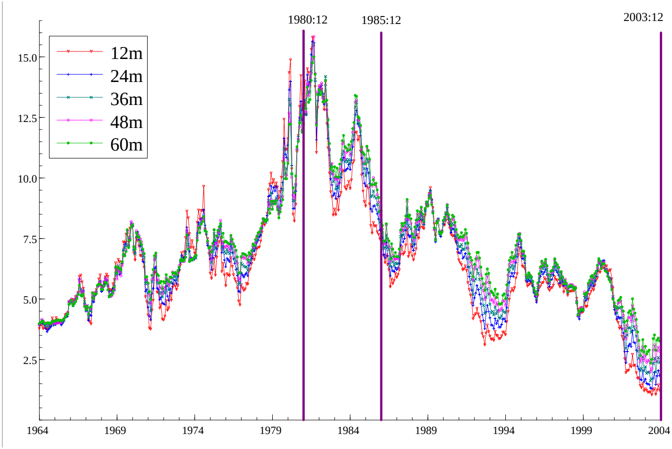
Draws, autocorrelations and posterior densities ϕ_h



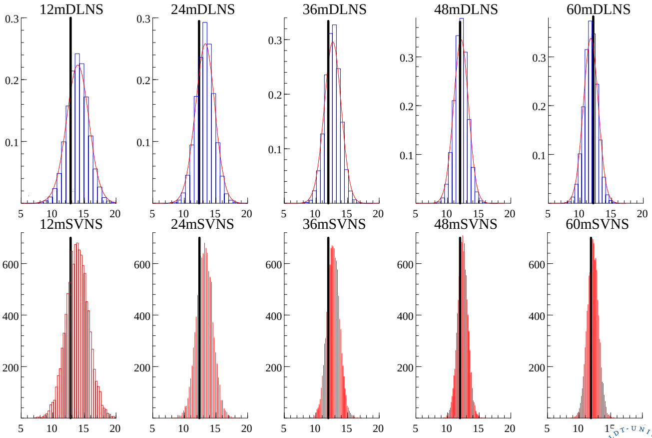
Draws, autocorrelations and posterior densities λ



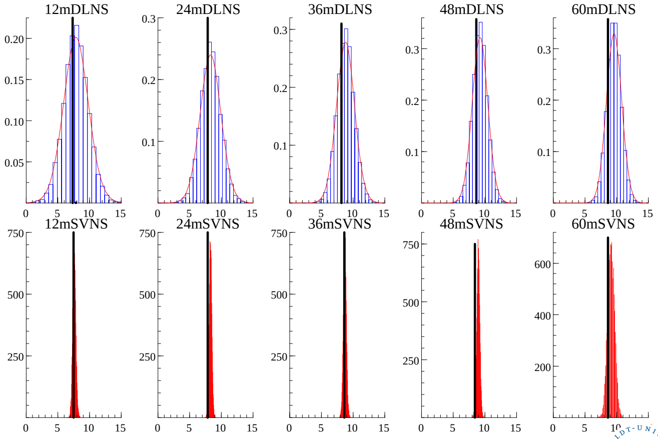
Forecast using 3 sub samples



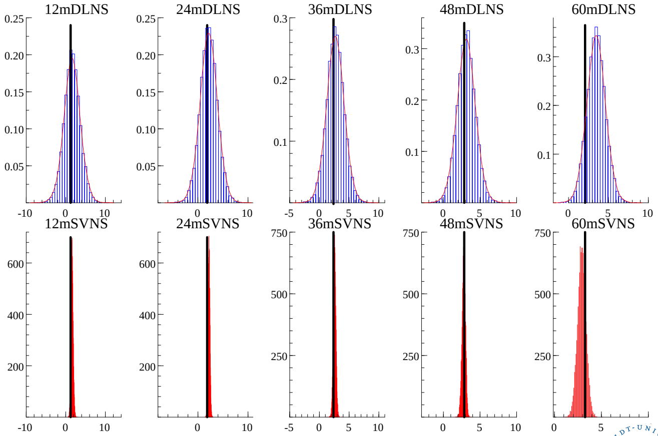
Forecast the yields for Dec, 1980



Forecast the yields for Dec, 1985



Forecast the yields for Dec, 2003



Outline

Conclusion

- The proposed algorithm is computationally efficient.
- The SVNS fits the data well according to posterior predictive p-value, see Hautsch and Yang (2010).
- Assuming a constant variance(i.e. DLNS model), the forecast density reveals larger uncertainty than it should be, since the data with high volatilities "contaminate" the estimate of the variances.
- Using Bayesian method, the uncertainty associates with the point forecast can be quantified.