Essays on Measuring Monetary Policy Uncertainty and Forecasting Business Cycle

By

Shangwen Huang

Submitted to the Department of Economics and the
Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Committee members

Shu Wu, Chairperson

John Keating

Iwata Shigeru

Xueming Tu

Jianbo Zhang

Date defended: April 29, 2016

The Dissertation Committee for Shangwen Huang certifies
that this is the approved version of the following dissertation :

Essays on Measuring Monetary Policy Uncertainty and Forecasting Business Cycle

Shu Wu, Chairperson

Date approved: May 26, 2016

Abstract

Chapter 1 is a survey paper for economists new to the field of a Monte Carlo simulation based method: particle filter. Particle filter can be applied to many flexible state space models such as non-linear, non-Gaussian, stochastic volatility models or stochastic volatility models with zero lower bound. These models have become increasingly popular in macro-economics and finance. The stochastic volatility model with zero lower bound is designed for applying zero interest rate policy. The primary purpose of this paper is to provide particle filter algorithms for these models and make comparisons of the Kalman filter, extended Kalman filter and particle filter.

In Chapter 2, we estimate a non-linear, non-Gaussian state space model for the short-term interest rate. The model features a potentially binding zero-lower bound constraint and stochastic volatility. We use the model to extract a measure of monetary policy uncertainty. We find that ignoring the zero-lower bound constraint on nominal interest rate leads to overestimation of the shadow rate and underestimation of the monetary policy uncertainty in recent periods. We find that the policy uncertainty is especially high at the beginning and the end of a quantitative easing episode. We also find a policy uncertainty shock lowers output growth and raises unemployment and seems to be a priced risk factor in the stock market.

In Chapter 3, I build a probit model of yields and business cycles. The proposed model incorporates the interrelation of yields and a latent business cycle factor, which will be extracted from the joint model. Different to previous literature, this model allows for interactions of the latent business factor and a portfolio of yields. The parameters and the latent variable are estimated through likelihood maximization at one step. I use the fully adapted particle filter to generate the likelihood of the nonlinear model. I find that the model with the autocorrelated latent variable forecasts better than the model without autocorrelated latent variable in terms of in-sample and out-of-sample forecast. The model implied business cycle factor indicates all 7 recessions from 1969 to 2015.

Acknowledgements

I would like to express my special appreciation and thanks to my advisor Professor Dr. Shu Wu, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been priceless. I would also like to thank my committee members, professor John Keating, professor Iwata Shigeru, professor Xueming Tu and professor Jianbo Zhang for serving as my committee members even at hardship. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

Contents

List of Figures

List of Tables

Chapter 1

Monte Carlo Study: Compare Particle Filter to Kalman Filter with Estimating Examples of Stochastic Volatility Model

1.1 Introduction

The relationships between economic variables are often nonlinear. In financial markets, time series exhibit non-Gaussian or time-varying volatility behavior. Recently the unconventional monetary policy in a zero interest rate environment arouses our interest in how to construct and solve a more realistic and flexible model. This paper surveys several state space models: a linear Gaussian model, a linear non-Gaussian model, and a nonlinear non-Gaussian model referring particularly to a stochastic volatility model (SV). This paper also constructs a stochastic volatility model with a zero lower bound (SV-ZLB). The SV-ZLB model would be more realistic for zero interest rate policy environment. Kalman filter (KF) provides an efficient solution on estimating linear Gaussian state space models, but it can not solve nonlinear or non-Gaussian models. In this paper, I introduce a Monte Carlo simulation based method: particle filter (PF) to solve non-liner, non-Gaussian state space models.

Extended Kalman filter (EKF) is based on KF to solve nonlinear models, but the model must be differentiable and Gaussian. Alternatively, PF can estimate a non-linear, non-Gaussian model. For EKF, people use Taylor expansion to linearize measurement equation by taking first order derivatives with respect to state variables; thus, the model must be differentiable. For particle filter, people use importance sampling to simulate latent state variables, a procedure, which does not have limitations on distribution or linearity. I will explain the theory of EKF and PF and further provide a PF algorithm: continuous sampling importance resampling (SIR) for estimating the SV model and SV-ZLB model.

First, I compare the estimation result of KF with PF for a linear Gaussian model and a linear Student's t model by using simulated data. Then, I compare the estimation result of EKF with PF for a SV model by using simulated data. In addition, I estimate the SV-ZLB model via PF with simulated data. The result shows both PF and KF can provide unbiased estimators and latent state variables for linear Gaussian model. PF has more computation burden than KF. PF is better than EFK on filtering states for linear non-Gaussian model and estimating transition equation for the SV model. Meanwhile, PF has no more computation burden than EFK. Finally, PF can provide accurate estimates and latent states for the SV-ZLB model.

The structure of the present paper is as follows. In section 2, I consider a SV model and construct a SV-ZLB model. I describe estimation methods in section 3, which includes the procedures of KF, EKF and PF. I describe continuous SIR algorithm for the SV and SV-ZLB model. In section 4, I compare the estimation results for linear Gaussian model, linear non-Gaussian model and SV model. Further, I estimate SV-ZLB model by PF with simulated data. The final section concludes.

1.2 Models

State space models are used to assume macroeconomic observations are determined by some underlying processes of state variables when solving economic problems. State space models contains two parts: measurement equation and transition equation, given by

$$
\begin{aligned}
y_t &= f(x_t, v_t) && (1.1)\\
x_t &= h(x_{t-1}, u_t) && (1.2)
\end{aligned}
$$

The measurement equation (1.1) describes how latent state variable x_t drives the movement of observations y_t. Transition equation (1.2) describes how state variables evolve over time. v_t and u_t denote observation noise and state noise respectively, which can be assumed in either Gaussian distribution or non-Gaussian distribution. f and h denote function forms that could be linear or nonlinear.

In Gaussian term structure models (GATSM), people generally assume homoskedasticity. Vasicek model is one of the standard short term interest rate models (Filipovic 2009, P85)

$$
dy = (b + \beta y)dt + \sigma dw \tag{1.3}
$$

where y represents an observed variable—interest rate. Equation (1.3) is a discrete-time approximation to a continuous-time Vasicek model given by (Harvey Ruiz and Shephard 1994)

$$
y_t = \rho_0 + \rho_1 y_{t-1} + \sigma u_t,\ u_t \sim N(0,1) \tag{1.4}
$$

Instead of assuming homoskedasticity, stochastic volatility (SV) models assume the logarithm of variance is in a stochastic process , accordingly $\sigma = e^{x_t/2}$ in (1.4)

$$
x_t = \phi x_{t-1} + w_t,\ w_t \sim N(0, \sigma^2) \tag{1.5}
$$

SV models allow for heteroskedasticity and let state variables determine observations' volatility.

I construct a new model: stochastic volatility model with zero lower bound (SV-ZLB) by combining the SV model and a Tobit model.

$$y_t \;=\; max\{\rho_0 + \rho_1 y_{t-1} + e^{x_t/2} u_t, 0\}, \; u_t \sim N(0,1) \tag{1.6}$$

x_t has the same process as (1.5). In (1.6), $\rho_0 + \rho_1 y_{t-1} + e^{x_t/2} u_t$ represent a shadow rate (Black 1995), which can drop freely below zero. The SV-ZLB model can capture the volatility of shadow rates when a nominal interest rate is stuck at zero lower bound.

1.3 Methods

1.3.1 Kalman filter and extended Kalman filter

Kalman filter (KF) is a widely used method for estimating linear state space models. KF delivers the true likelihood, thus can be used to estimate parameters via maximizing likelihood (MLE). KF algorithm contains predicting steps and updating steps. On predicting steps, the expected value of state variable is calculated at t given all the information known up to $t-1$. In updating steps, by taking into account Kalman gain, which is the weight put on the new observation (Kalman 1960, Kalman and Bucy 1961), the expected value of state variables would be updated conditional on information at t. Through predicting and updating steps recursively, the predicted and updated states would be obtained for each period time. For linear Gaussian state space models, the log-likelihood can be calculated with the prediction error and its variance, see (Kim and Nelson 1999). For linear non-Gaussian state space model, the quasi log-likelihood would be calculated as Gaussian.

Extended Kalman filter (EKF) has been developed for non-linear, non-Gaussian state space model. For instance, the SV model is a non-linear and non-Gaussian model, thus can not be estimated by KF. As an extension to KF, EKF utilize the Laplace approximation as an asymptotic approximation to the posterior distribution (Gamerman 1997). Here I follow the notations in (Mayer at al. 2003). First $p(x_t|Y_{t-1})$ is approximated by a normal PDF $\tilde{p}(x_t|Y_{t-1})$ with mean: $b(t) = \phi\hat{x}_{t-1}$ and variance: $\gamma_t^2 = \phi^2\hat{s}_{t-1}^2 + \sigma^2$, where Y_{t-1} denotes $\{y_1, y_2......y_{t-1}\}$. Then the likelihood (7) can be approximated by using Laplace approximation (Meyer at al. 2003) as (1.8).

$$
\begin{aligned}
p(y_t|Y_{t-1},\theta) &\approx \int p(y_t|x_t,\theta)\tilde{p}(x_t|Y_{t-1},\theta)dx_t && (1.7)\\
&= \sqrt{2\pi}e^{-\psi_t(y_t,\hat{x}_t,\theta)}|D^2\psi_t(y_t,\hat{x}_t,\theta)|^{-1/2} && (1.8)
\end{aligned}
$$

where

$$\psi_t(y_t,x_t,\theta) = -log(p(y_t|x_t,\theta)\tilde{p}(x_t|Y_{t-1},\theta)) \quad (1.9)$$

$$\tilde{p}(Y|\theta) = p(y_1|\theta)\prod_{t=2}^{T} p(y_t|Y_{t-1},\theta) \quad (1.10)$$

Where, $\hat{x}_t = argmin_{x_t}\psi_t(y_t,x_t,\theta)$, $\hat{s}_t^2 = |D^2\psi_t(y_t,\hat{x}_t,\theta)|^{-1}$, $D^2\psi_t(y_t,\hat{x}_t,\theta)$ denotes the second-order derivative of the function $\psi_t(y_t,x_t,\theta)$ with respect to x_t. Finally, once the likelihood (1.10) is obtained, the parameters can be estimated by MLE. In this chapter, I use the EKF to estimate the parameters for the SV model. Because the latent state variables are integrated out in (1.7), I provide the state variables by predicted mean $b(t)$.

1.3.2 Particle filter

Alternatively, particle filter (PF) is a simulation-based technique for estimating non-linear, non-Gaussian state space models, also called Sequential Monte Carlo. First, I describe how to extract latent state variables through PF by assuming the parameters are fixed and then consider parameter estimation in the latter paragraphs. Let $x_{0:t}$ denote $\{x_0, x_1, ..., x_t\}$ and $y_{1:t}$ denote $\{y_1, y_2, ..., y_t\}$. For any function f, the conditional expectation of state variables is given by

$$E[f(x_{0:t})|y_{1:t}] = \int f(x_{0:t})p(x_{0:t}|y_{1:t};\theta)dx_{0:t} \quad (1.11)$$

where the target distribution $p(x_{0:t}|y_{1:t};\theta)$ is referred to the posterior PDF, which is non-standard and unknown (Doucet 2001). The principle of Monte Carlo is that latent state variables can be obtained through simulations. For PF, simulations are completed by drawing particles x_t^i from a proposal distribution $g_{0:t}(x_{0:t}|y_{1:t};\phi)$. Let i denote the ith particle.

However, the target cannot be the same as proposal, discrepancy exists between the two distributions. The discrepancy is measured by an importance weight, which is calculated by taking target density divided by proposal density(Creal 2009) as (1.12). Each particle x_t^i will have it's own weight w_t^i.

$$w_t^i \propto \frac{p(x_{0:t}^i|y_{1:t};\theta)}{g_{0:t}(x_{0:t}^i|y_{1:t};\phi)} \tag{1.12}$$

The definition of importance weight implies that if one particle is applied with a high proposal density but actually has a low target density, the particle will be corrected by applied with a small weight.

$$E[f(x_{0:t})|y_{1:t}] = \int \frac{f(x_{0:t})p(x_{0:t}|y_{1:t};\theta)}{g_{0:t}(x_{0:t}|y_{1:t};\phi)} g_{0:t}(x_{0:t}|y_{1:t};\phi)dx_{0:t} \tag{1.13}$$

$$= \int f(x_{0:t})w_t^i g_{0:t}(x_{0:t}|y_{1:t};\phi)dx_{0:t} \tag{1.14}$$

By multiplying and dividing the importance distribution in the integral (1.11), the conditional expectation of latent state variables becomes an integral of all weighted particles in (1.14).

In discrete form, (1.11) is approximated by the weighted summation of particles

$$E[f(x_{0:t})|y_{1:t}] \approx \sum_{i=1}^{N} f(x_{0:t}^i)\hat{w}_t^i \tag{1.15}$$

$$\hat{w}_t^i = \frac{w_t^i}{\sum_{i=1}^{N} w_t^i} \tag{1.16}$$

where, $\hat{w}_t^i$ is the normalized weight for particle i. Furthermore, apply Baye's rule to (1.14)

$$g_{0:t}(x_{0:t}|y_{1:t};\phi) \equiv g_t(x_t|x_{0:t-1},y_{1:t};\phi)g_{0:t-1}(x_{0:t-1}|y_{1:t-1};\phi) \tag{1.17}$$

$$w_t^i \propto \frac{p(y_t|x_t;\theta)p(x_t|x_{t-1};\theta)}{p(y_t|y_{1:t};\theta)g_t(x_t|x_{0:t-1},y_{1:t};\phi)}\frac{p(x_{0:t-1}|y_{1:t-1};\theta)}{g_{0:t-1}(x_{0:t-1}|y_{1:t-1};\phi)} \tag{1.18}$$

Suppose state variables are in Marcov process, the weight results in the product of incremental weight and previous weight:

$$w_t^i \quad \propto \quad \frac{p(y_t|x_t;\theta)p(x_t|x_{t-1};\theta)}{g_t(x_t|x_{t-1},y_{1:t};\phi)}w_{t-1}^i \tag{1.19}$$

and in (1.17) the incremental importance weight is defined as

$$\tilde{w}_t^i \quad \equiv \quad \frac{p(y_t|x_t;\theta)p(x_t|x_{t-1};\theta)}{g_t(x_t|x_{t-1},y_{1:t};\phi)} \tag{1.20}$$

Bootstrap particle filter assumes $p(x_t|x_{t-1};\theta) = g_t(x_t|x_{t-1},y_{1:t};\phi)$, then the incremental weight $p(y_t|x_t;\theta)$ is referred to the likelihood of observation conditional on the drawn particles. By given the initial particles $\{x_0^i\}_{i=1:N}$ and their weights $\{w_0^i\}_{i=1:N}$, recursively state variables can be extracted in each time period.

1.3.3 Sampling Importance Resampling (SIR)

Sampling importance resampling (SIR) is a bootstrap PF algorithm. This chapter uses SIR to estimate state space models. The SIR algorithm is a step-by-step procedure of how to draw particles and calculate importance weights over time. Rubin (1987, 1988) first proposed the SIR algorithm.

SIR algorithm: Given the particles x_t^i for i=1,...,N. At time t:

(i) Sample x_t^i from $p(x_t|x_{t-1}^i)$

(ii) Compute the normalized importance weight of x_t^i:

$$w_t^i = \frac{p(y_t|x_t^i)}{\sum_{i=1}^N p(y_t|x_t^i)} \tag{1.21}$$

(iii) Resample with replacement: draw $k(i) \sim Mult_N(\{w_t^i\}_{i=1}^N)$

(iv) Set $x_t^i = x_t^{k(i)}$, $\{x_t^i, \frac{1}{N}\}$

The SIR algorithm involves drawing particles x_t^i from $p(x_t|x_{t-1}^i)$, resampling $x_t^k(i)$ from weights w_t^i and applying new particles with the same weight $1/N$. The expected value of state variables would be the sample average.

Without resampling, equation (1.17) and (1.18) show particles' weight accumulate over time, so some particles with big weight would have bigger weight and some particles with small weight would have smaller weight. All the probability mass eventually would be allocated to one particle as the number of iterations increases, that phenomenon is called the degenerate problem (Chopin 2004). With resampling, a smaller weight particle will become a smaller numbers of particles and a bigger weight particle will become a bigger numbers of particles, but all of the particles have the equalized weight. So, weights no longer accumulate but convert to numbers of particles. Thus, the SIR algorithm can help to alleviate the degenerate problem from not losing some of the particles.

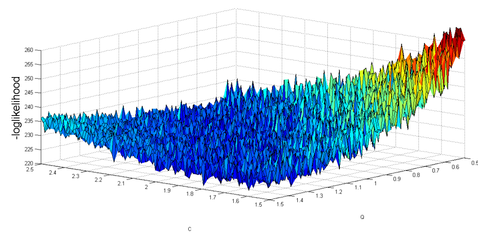


Figure 1.1: 2D discontinuous negative log likelihood

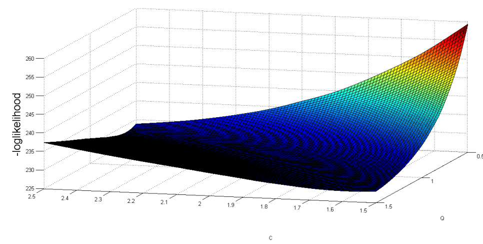


Figure 1.2: 2D smoothed negative log likelihood

1.3.3.1 Likelihood function

In this paper, parameters are estimated by assuming the initial value of parameters and searching maximum likelihood by Newton-Rhapson methods. The likelihood function is given by

$$\begin{aligned}
\mathscr{L}(\theta|y_{1:T})) &= \prod_{t=2}^{T} p(y_t|\theta, y_{1:t-1}) && (1.22)\\
&= \prod_{t=2}^{T} \int p(y_t|x_t,\theta)p(x_t|y_{1:t-1},\theta)dx_t && (1.23)\\
&= \prod_{t=2}^{T} \frac{1}{N} \sum_{i=1}^{N} w_t^i && (1.24)
\end{aligned}$$

In equation (1.24), the likelihood function is the product of averaged weight. Since, the density $p(x_t|x_{t-1}^i)$ and $p(y_t|x_t^i)$ are assumed in the state space model, by going through step (i) to step (iv). Then, importance weight will be obtained and estimators can be estimated by maximizing likelihood in (1.21).

A problem by using MLE is the resampling step causes discontinuities in the estimated likelihood with respect to parameters. Likelihood results in a rough surface, thus to find global maximum is not easy. Figure 1 shows a 2-dimension negative log likelihood. (Malik and Pitt 2011) provide a continuous resampling procedure to smooth the likelihood. It approximates linear interpolation for Gaussian samples and have a set of uniforms to invert this approximated CDF. Particles need to be sorted in ascending order before and after resampling step for each time period. By following the continuous resampling procedure, a 2-dimension negative log likelihood will be like Figure 1.2. Then MLE can be applied to estimate parameters. Del Moral (2004) has demonstrated that the particle filter delivers a consistent and an unbiased estimator for the true likelihood function. $\hat{L}_N(\theta) \xrightarrow{a.s} L(\theta)$ as $N \to \inf$ and $E[\hat{L}_N(\theta)] = L(\theta)$.

1.4 Examples: Compare KF, EKF to SIR

1.4.1 Linear Gaussian Model

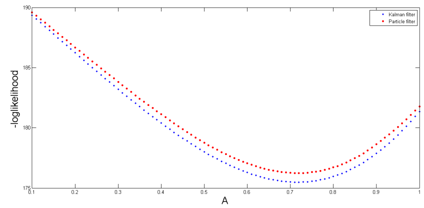
To find the difference in estimation performance of Kalman filter and particle filter, I first consider a linear Gaussian state space model. Let y_t denote observation and x_t denote the latent state variable. Both observation noise u_t and state noise w_t follow a Gaussian distribution with variance R and Q respectively.

$$y_t = Cx_t + u_t, \; u_t \sim N(0, R) \tag{1.25}$$

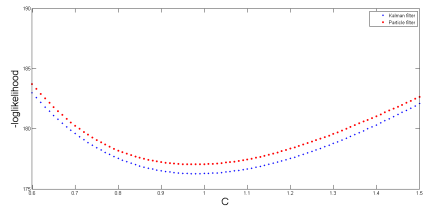
$$x_t = Ax_{t-1} + w_t, \; w_t \sim N(0, Q) \tag{1.26}$$

To investigate the effects of length of data and number of particles on estimation performance. I simulate two groups of data $T = 100$ and $T = 300$. For particle filter, 80 and 150 particles are used for two groups data respectively. True values of parameter are: $A = 0.6$, $C = 1$, $Q = 1$ and $R = 1$.

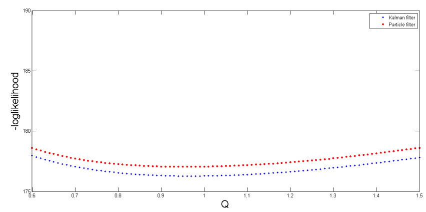
In figure 1.3, I range A from 0.1 to 1 while fixing other parameters at the true values and range C, Q and R from 0.5 to 1.5 respectively while fixing other parameters at the true values to obtain the sliced negative log-likelihood by Kalman filter and by particle filter. In each graph, Kalman



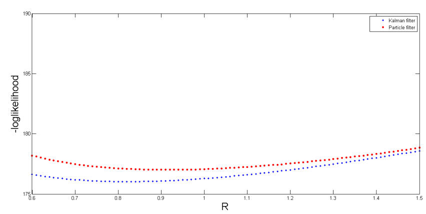
(a) Varying values for parameter A



(b) Varying values for parameter C



(c) Varying values for parameter Q



(d) Varying values for parameter R

Figure 1.3: Sliced negative log likelihood by Kalman filter (blue) and by particle filter (red) using simulated data. $N = 80$, $T = 100$, true value $A = 0.6, C = 1, Q = 1, R = 1$

filter likelihood is larger than particle filter likelihood, because Kalman filter likelihood is the exact likelihood and particle filter likelihood is a approximated one for the linear Gaussian state space model. In the (a) and (b), the distance between approximated likelihood and true likelihood increases when approaching to peak and decreases toward two ends. Comparatively, the likelihoods for R and Q are much flatter. The maximum likelihood are obtained at almost the same value of parameters via Kalman filter and particle filter (ie. $A = 0.73$ and $C = 0.98$).

	estimation	grad $\times 10^6$	var$\times 10^3$	
A	0.7805	-3.81	8.29	-7.79
C	0.8184	3.81	-7.79	22.29
LL	-174.83			

(a) $T = 100$

	estimation	grad $\times 10^5$	var$\times 10^3$	
A	0.5838	-3.05	5.25	-3.13
C	0.9814	-3.81	-3.13	7.64
LL	-535.72			

(b) $T = 300$

Table 1.1: Estimation results by Kalman filter for linear Gaussian model by simulated data. $T = 100$ and $T = 300$ respectively, true value: $A = 0.6, C = 1, Q = 1, R = 1$. Fix Q and R at true value.

Table 1.1 shows the estimation results for the linear Gaussian state space model by using Kalman filter with fixing Q and R at the true value. The variance is obtained by taken inverse of the negative log-likelihood's Hessian. The initial values of parameters we set are $A_0 = 0.5$, $C_0 = 0.9$. Table 1a uses data with length $T = 100$ and table 1b uses data with length $T = 300$. In 1a, $\hat{A} = 0.7805$, $\hat{C} = 0.8184$ and in 1b, $\hat{A} = 0.5838$, $\hat{C} = 0.9814$. The estimate for A in table 1a is 30% higher than the true value and by in table 1b is 3% lower than the true value. The estimate for C in table 1a is 20% lower than the true value and in table 1b is 2% lower than the true value. The estimates will be more accurate with longer time series land larger numbers of particles. Kalman filter provide unbiased estimates for the linear Gaussian state space model when simulated data length is 300.

Table 1.2 shows the estimation results for the linear Gaussian state space model by Particle filter with fixing Q and R at the true value. Table 1.2a uses data with length $T = 100$ and 80 particles and table 2b uses data with length $T = 300$ and 150 particles. Let the initial values be $A_0 = 0.5$ and $C_0 = 0.5$. While running MLE with continuous SIR, seeds must be fixed. To show whether setting different seeds would affect estimates, I list the estimates by using 5 seeds and the average of 5 seeds' estimates.

Now we look at table 1.2 to examine Particle filter's estimates by different seed. In table 1.2b, estimates by 5 seeds are not as variant as in table 1.2a. In table 1.2a, the result varies from the best

seed1	est	grad$\times10^2$	var$\times10^2$		seed2	est	grad$\times10^2$	var$\times10^2$	
A	0.7722	9.38	0.12	0.01	A	0.7350	9.46	0.20	0.11
C	0.8496	-1.88	0.01	0.20	C	0.8424	-4.07	0.11	0.33
LL	-175.57				LL	-175.07			
seed3	est	grad$\times10^2$	var$\times10^2$		seed4	est	grad$\times10^2$	var$\times10^2$	
A	0.7561	4.37	0.05	0.00	A	0.8008	1.35	0.05	0.07
C	0.8485	5.08	0.00	0.09	C	0.8005	-0.12	0.07	0.64
LL	-175.34				LL	-175.48			
seed5	est	grad$\times10^2$	var$\times10^2$		avg	est	grad$\times10^2$	var$\times10^2$	
A	0.8324	2.12	0.04	-0.04	A	0.7793	5.33	0.09	0.03
C	0.7892	0.47	-0.04	0.17	C	0.8260	-0.10	0.03	0.29
LL	-175.22				LL	-175.34			

(a) $N = 80, T = 100$

seed1	est	grad$\times10^2$	var$\times10^2$		seed2	est	grad$\times10^2$	var$\times10^2$	
A	0.5641	10.80	0.13	0.18	A	0.5641	1.89	0.28	-0.21
C	0.9972	-2.28	0.18	0.60	C	0.9891	-1.19	-0.21	0.49
LL	-536.69				LL	-533.95			
seed3	est	grad$\times10^2$	var$\times10^2$		seed4	est	grad$\times10^2$	var$\times10^2$	
A	0.5810	0.85	0.35	-0.32	A	0.5929	0.84	0.08	-0.02
C	0.9884	0.18	-0.32	0.66	C	0.9875	0.42	-0.02	0.20
LL	-537.37				LL	-536.98			
seed5	est	grad$\times10^2$	var$\times10^2$		avg	est	grad$\times10^2$	var$\times10^2$	
A	0.5714	2.30	0.14	-0.08	A	0.5747	3.33	0.20	-0.09
C	0.9712	-0.42	-0.08	0.20	C	0.9867	-0.66	-0.09	0.43
LL	-536.45				LL	-536.29			

(b) N=150,T=300

Table 1.2: Estimation results by particle filter with different seeds for linear Gaussian model. True value: $A = 0.6, C = 1, Q = 1, R = 1$. Fix Q and R ate true value.

one, which is given by seed 2 $\hat{A} = 0.7350$, $\hat{C} = 0.8424$ to the worse one, which is given by seed 5 $\hat{A} = 0.8324$, $\hat{C} = .7892$. In table 1.2b, the best one is from seed 4 $\hat{A} = 0.5929$, $\hat{C} = 0.9875$ and the average are $\hat{A} = 0.5747$, $\hat{C} = 0.9867$. Neither the worst or the best is too far from the average when using data length $T = 300$. Therefore, if the sample time series is long enough, estimates by different seeds would have little standard deviation and closer to true value.

In table 1.2a, the average estimates are $\hat{A} = 0.7793$, $\hat{C} = 0.8260$. In table 1.2b, the average estimates are $\hat{A} = 0.5747$, $\hat{C} = 0.9867$. When $T = 100$, the estimate of A is 29.9% higher than the true value and the estimate of C is 18% lower than the true value. When $T = 300$, the estimate of A is 5% lower than the true value and of C is 2% lower than the true value. For particle filter, the estimates are more accurate when using longer time series. Particle filter provides accurate estimates when we using data with length of $T = 300$.

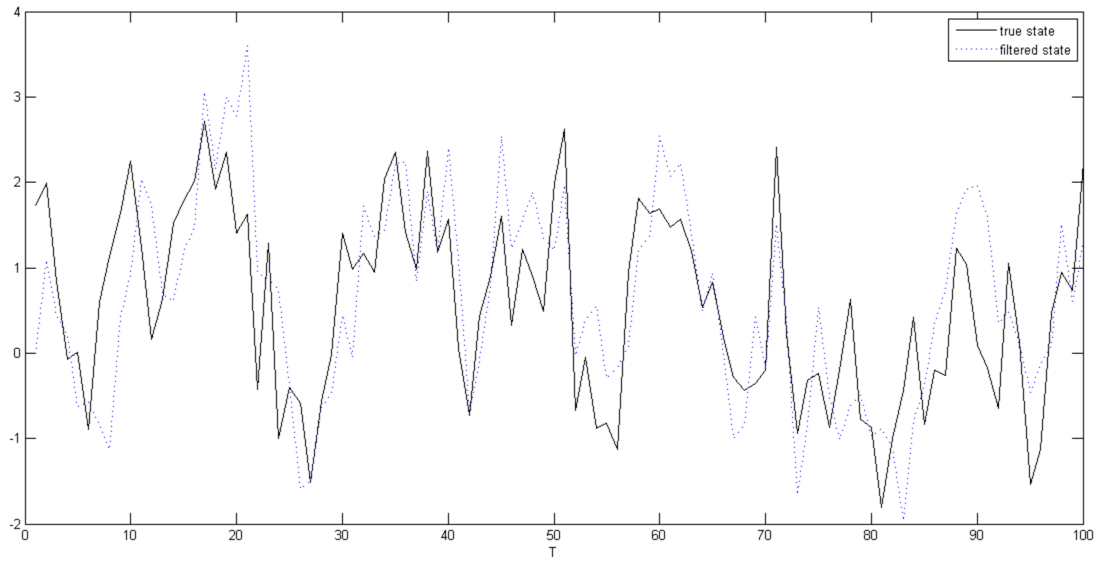
Then, let us compare particle filter to Kalman filter on estimating the linear gaussian state space model. For $T = 300$ case, Particle filter's average estimates are $\hat{A} = 0.5747$, $\hat{C} = 0.9867$, compared to Kalman filter's estimates, which are $\hat{A} = 0.5838$, $\hat{C} = 0.9814$. Figure 4 and 5 are the filtered state vs. true state by Kalman filter and particle filter for $T = 100$ and $T = 300$ respectively. The solid line is real state from simulated data and the dotted line is filtered state. Based on our experiment, particle filter's estimates are as good as Kalman filter. And both Kalman filter and particle filter can extract latent state variables accurately as the length of time series is long enough.

1.4.2 Linear non-Gaussian model

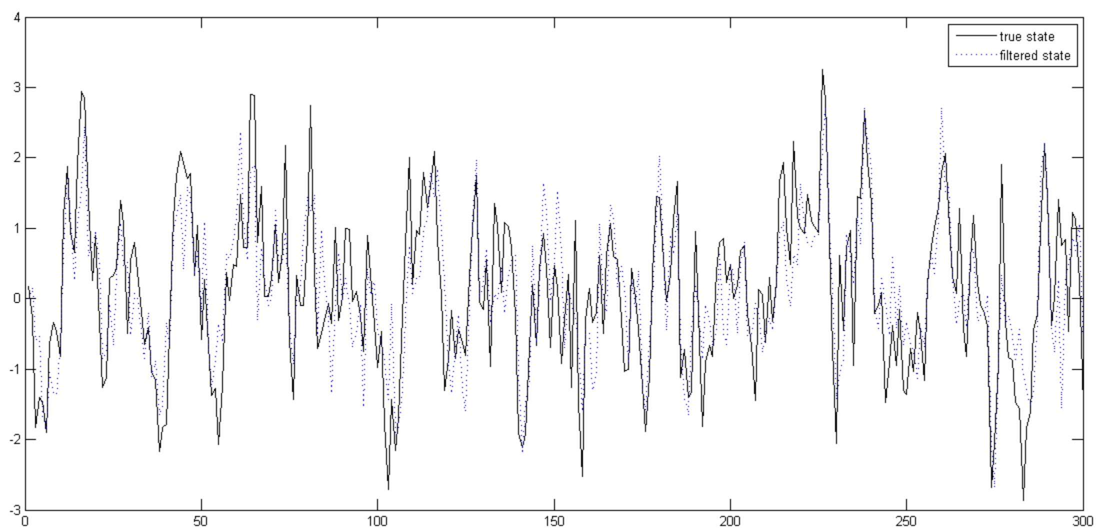
- To compare the estimation performance of Kalman filter and particle filter for linear non-Gaussian state space model, I assume observation noise follows Student's t distribution. V is the degree of freedom of t distribution and variance is defined as $\frac{V}{V-2}$.

$$y_t = Cx_t + u_t, \ u_t \sim t(V) \tag{1.27}$$

$$x_t = Ax_{t-1} + w_t, \ w_t \sim N(0, Q) \tag{1.28}$$

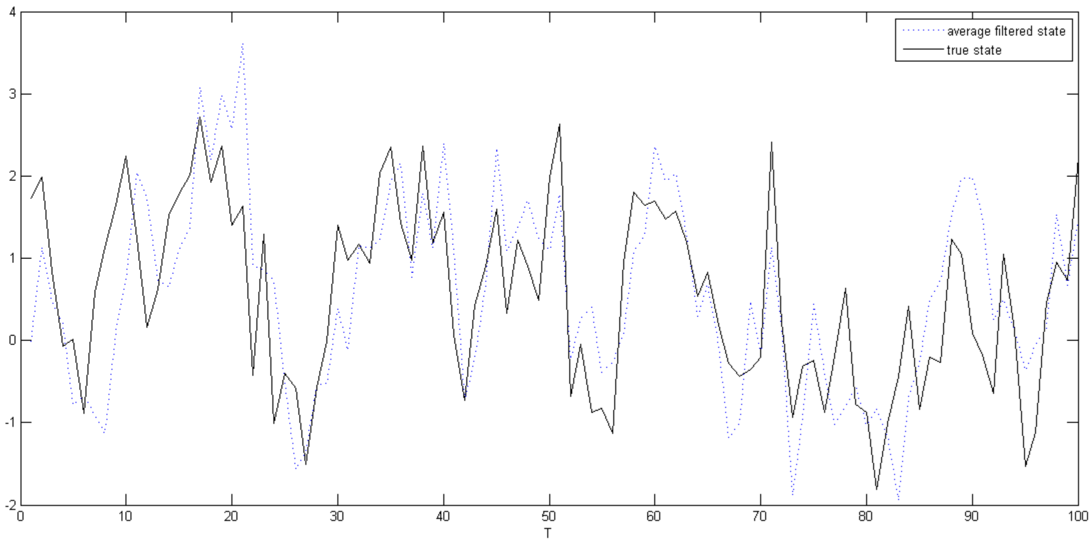


(a) $T = 100$

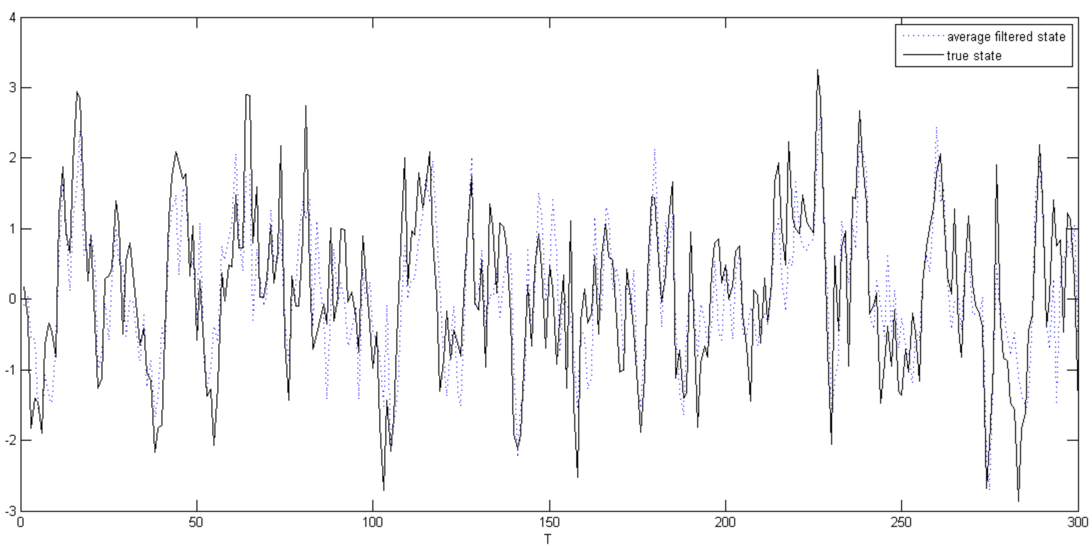


(b) $T = 300$

Figure 1.4: Filtered state by Kalman filter for linear Gaussian model. $T = 100$ and $T = 300$ respectively, true value: $A = 0.6, C = 1, Q = 1, R = 1$. Fix Q and R at true value. Solid line is true state and dashed line is filtered state.



(a) $N = 80, T = 100$



(b) $N = 150, T = 300$

Figure 1.5: Filtered state by particle filter for linear Gaussian model. Use $N = 80, T = 100$ and $N = 150, T = 300$ respectively, true value: $A = 0.6, C = 1, Q = 1, R = 1$. Fix Q and R at true value. Solid line is true state and dashed line is particles averaged filtered state.

	estimation	grad$\times 10^6$	var$\times 10^2$	
A	0.6807	-7.63	1.04	-1.21
C	0.9814	0.00	-1.21	3.34
$-LL$	429.56			

Table 1.3: Estimation results by quasi likelihood of Kalman filter for linear student's t model. $T = 200$, true value: $A = 0.6, C = 1, Q = 1, R = 3$(variance of t distribution when degree of freedom is 3). Fix Q an dR at true value.

seed1	est	grad$\times 10^2$	var$\times 10^2$		seed2	est	grad$\times 10^2$	var$\times 10^2$	
A	0.6372	-0.25	0.11	0.13	A	0.6552	-4.26	0.09	-0.05
C	1.0718	0.17	0.13	0.88	C	1.0614	4.54	-0.05	0.11
$-LL$	424.00				$-LL$	533.95			
seed3	est	grad$\times 10^2$	var$\times 10^2$		seed4	est	grad$\times 10^2$	var$\times 10^2$	
A	0.6016	2.07	0.59	-0.48	A	0.6489	24.51	0.04	-0.02
C	1.2204	1.01	-0.48	0.55	C	1.1550	2.79	-0.02	0.18
$-LL$	421.3615				$-LL$	425.0776			
seed5	est	grad$\times 10^2$	var$\times 10^2$		avg	est	grad$\times 10^2$	var$\times 10^2$	
A	0.5926	0.70	0.09	0.01	A	0.6271	4.55	0.18	-0.08
C	1.1749	0.34	0.01	0.12	C	1.1367	1.77	-0.08	0.37
$-LL$	421.13				$-LL$	445.10			

Table 1.4: Estimation results by particle filter for linear student's t model. $T = 200$, true value:$A = 0.6, C = 1, Q = 1, V = 3$ (degree of freedom is 3). Fix Q and R at true value.

The simulated data has $T = 200$, true value $A = 0.6, C = 1, Q = 1, V = 3$. Although Kalman filter can not provide true likelihood for non-Gaussian model, I use quasi likelihood by treating observation noise as a Gaussian model with variance 3.

Table 1.3 and table 1.4 show estimation result of Kalman filter and particle filter with 5 different seeds. Kalman filter estimators are $\hat{A} = 0.6807$, $\hat{C} = 0.9917$. Particle filter averaged estimators by 5 seeds are $\hat{A} = 0.6$, $\hat{C} = 1.1367$. For A, Kalman filter estimate is 13.45% higher than the true value and particle filter estimate is 4.5% higher than true value. For C, Kalman filter estimate is 0.83% lower than the true value and particle filter estimate is 13.7% higher than the true value. By all seeds except for seed 2, particle filter provides a higher log likelihood than Kalman filter.

Figure 1.6 and figure 1.7 present the Kalman filtered state and particle filtered state with true state respectively. The solid lines are true value and the dotted lines are model estimated state.

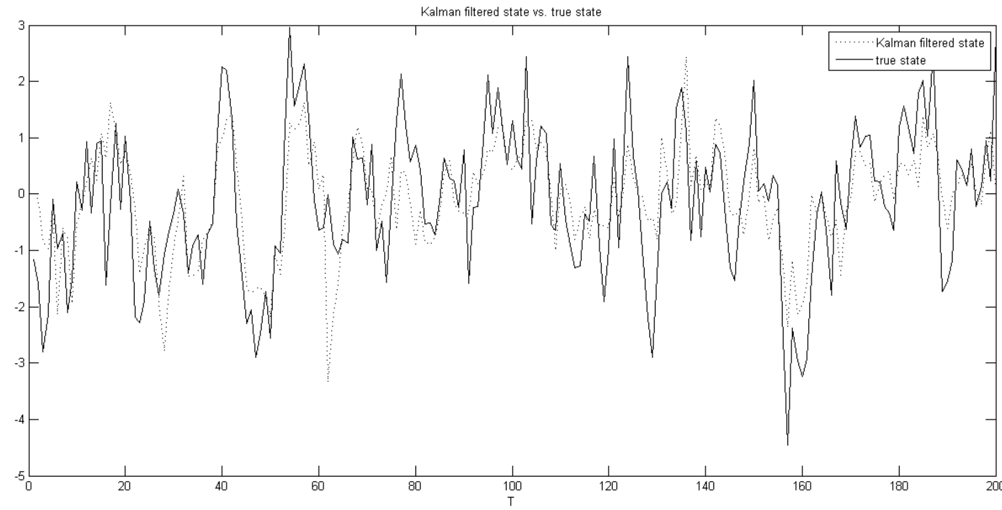


Figure 1.6: Filtered state by Kalman filter for linear studen's t model. $T = 200$, true value: $A = 0.6, C = 1, Q = 1, R = 3$. Fix Q and R at true value. Solid line is true state and dashed line is Kalman filtered state.

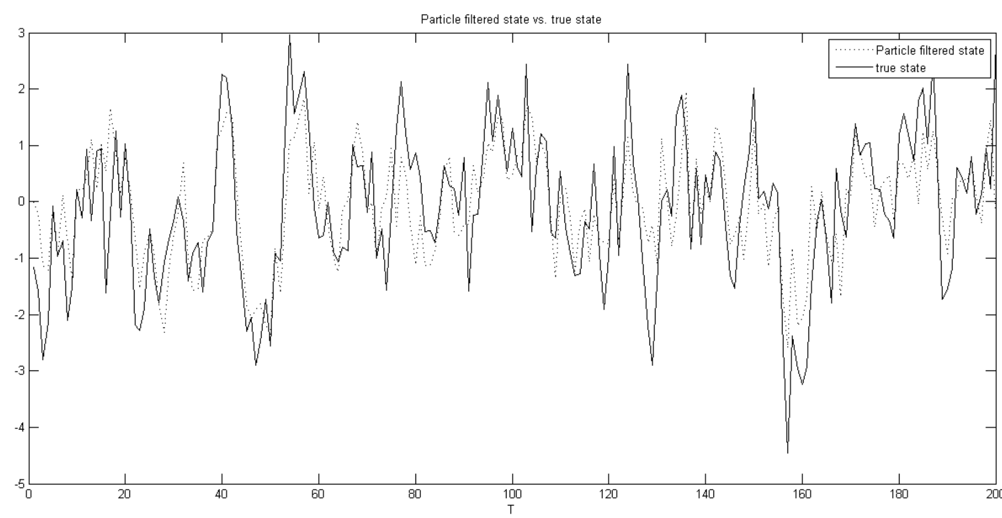


Figure 1.7: Filtered state by particle filter for linear studen's t model. $N = 100, T = 200$, true value: $A = 0.6, C = 1, Q = 1, V = 3$. Fix Q and V at true value. Solid line is true state and dashed line is averaged particle filtered state.

	mean	std err
ρ_0	-0.0020	0.8367
ρ_1	0.9516	0.030
ϕ	0.6929	0.1345
σ^2	0.8881	0.0819
LL	-372.07	67.97

Table 1.5: Extended Kalman filter bootstrap with 200 replicates for stochastic volatility model. $T = 200$, true value: $\rho_0 = 0, \rho_1 = 0.9, \phi = 0.9, \sigma^2 = 1$.

	mean	std err
ρ_0	-0.0040	0.0632
ρ_1	0.9544	0.0141
ϕ	0.8871	0.0412
σ^2	1.1145	0.2919
LL	-320.84	61.07

Table 1.6: Particle filter bootstrap with 200 replicates for stochastic volatility model. $N = 100, T = 200$, true value: $\rho_0 = 0, \rho_1 = 0.9, \phi = 0.9, \sigma^2 = 1$.

Both Kalman filtered state and particle filtered state are close to the true state except for $t = 60$ where Kalman filtered state is around three times lower than the true state.

1.4.3 Stochastic volatility model

Stochastic volatility state space model is as equation (4) (5). Here I rewrite the model:

$$y_t = \rho_0 + \rho_1 y_{t-1} + e^{x_t/2} u_t, \; u_t \sim N(0,1) \tag{1.29}$$

$$x_t = \phi x_{t-1} + w_t, \; w_t \sim N(0, \sigma^2) \tag{1.30}$$

Table 1.5 and table 1.6 are SV model estimated results by EKF and PF respectively with 200 replications of simulated data by changing noise. The length of each simulated data is 200. The true values of parameters for simulated data are: $\rho_0 = 0$, $\rho_1 = 0.9$, $\phi = 0.9$ and $\sigma^2 = 1$.

In table 1.5, the mean of $\hat{\phi}$ estimated by EKF is 0.6929 and standard error is 0.1345. In table 1.6, the mean of $\hat{\phi}$ estimated by PF is 0.8871 and the variance is 0.0412. EKF estimator for ϕ is

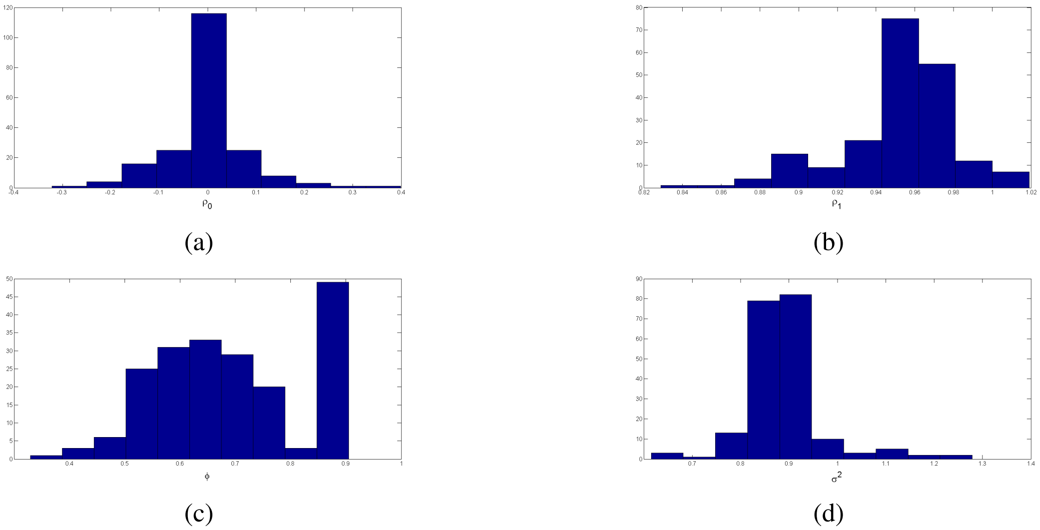


Figure 1.8: Histogram of estimates for $\rho_0, \rho_1, \phi, \sigma^2$ by 200 bootstrapping of extended Kalman filter for stochastic volatility model. $T = 200$, true value$\rho_0 = 0, \rho_1 = 0.9, \phi = 0.9, \sigma^2 = 1$.

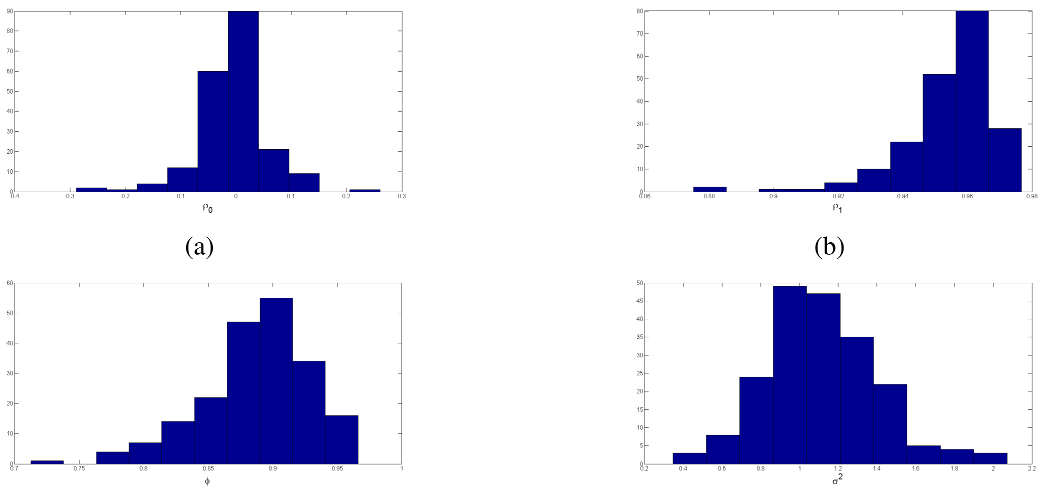
23% lower than the true value and PF esimator is 1.4% lower than the true value. Therefore, when estimating the SV model, PF has a more accurate estimator for volatility process than EKF. And PF estimated log likelihood is larger than EKF log likelihood by 50. The histogram of estimated parameters by EKF and PF are figure 8 and figure 9 respectively. Compared 8(c) to 9(c), ϕ's histogram of PF are more look like a normal distribution, but of EKF has many numbers allocate on the right end. Figure 10 shows estimated state of EKF and PF vs. true state by using the mean of estimates in table 5 and 6 respectively. The two filtered states do not have much differences and both are close to true state.

1.4.4 Stochastic volatility model with zero lower bound

Stochastic volatility model with zero lower bound is as equation (1.5) (1.6), here I rewrite the model

$$y_t = max\{\rho_0 + \rho_1 y_{t-1} + e^{x_t/2} u_t, 0\}, \; u_t \sim N(0,1) \tag{1.31}$$

$$x_t = \phi x_{t-1} + w_t, \; w_t \sim N(0, \sigma^2) \tag{1.32}$$



(a) (b)

(c) (d)

Figure 1.9: Histogram of estimates for $\rho_0, \rho_1, \phi, \sigma^2$ by 200 bootstrapping of particle filter for stochastic volatility model. $N = 100, T = 200$, true value$\rho_0 = 0, \rho_1 = 0.9, \phi = 0.9, \sigma^2 = 1$.

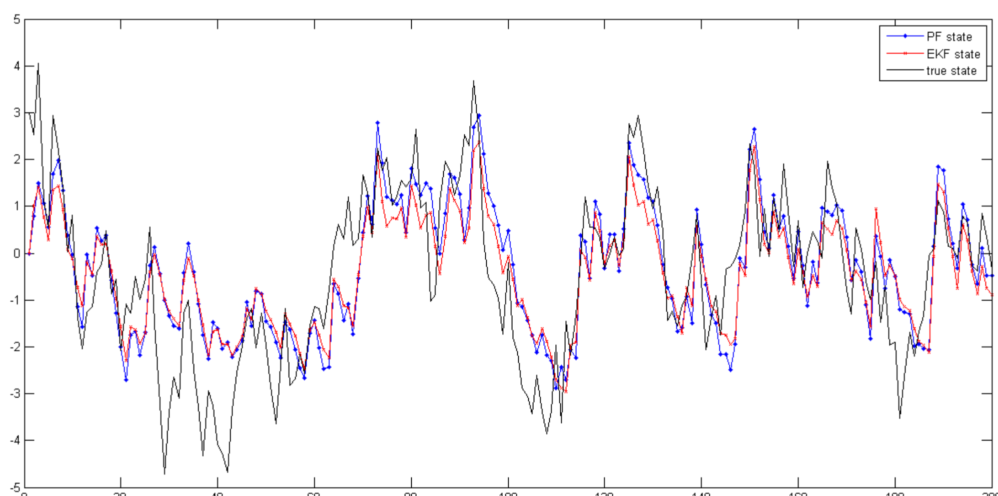


Figure 1.10: Extend Kalman filtered state, particle filtered state vs. true state for stochastic volatility model.

	est	grad	var $\times 10^4$			
ρ_0	0.0539	0.5805	1.97	0.01	0.34	1.35
ρ_1	0.9570	-1.3116	0.01	0.08	0.01	0.13
ϕ	0.9180	-0.6629	0.34	0.01	0.63	-0.30
σ^2	0.9636	-0.0995	1.35	0.13	-0.20	7.97
$-LL$	297.47					

Table 1.7: Estimation results by particle filter for stochastic volatility model with zero lower bound. $N = 100, T = 200$, true value: $\rho_0 = 0, \rho_1 = 0.96, \phi = 0.9, \sigma^2 = 1$.

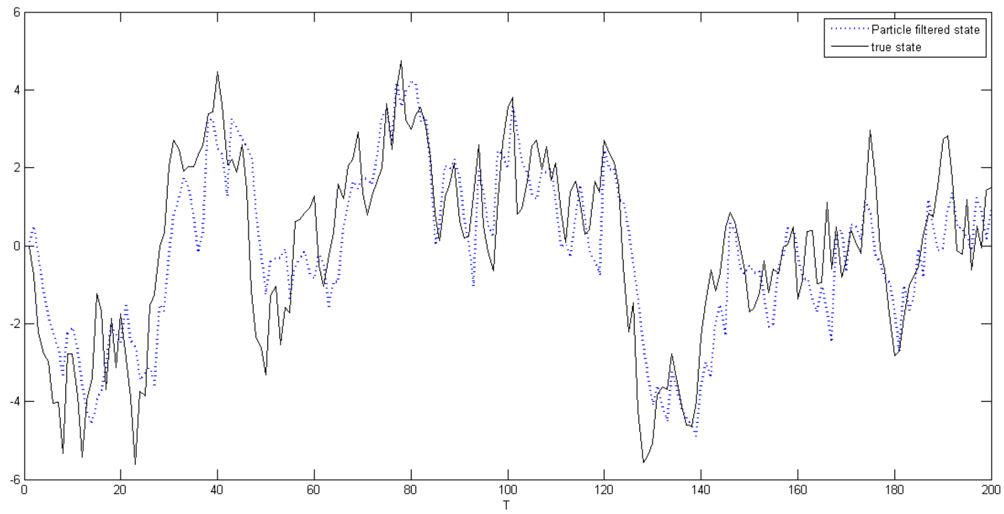


Figure 1.11: Particle filtered state vs. true state for stochastic volatility model with zero lower bound.

Because SV-ZLB model is nonlinear and non-differentiable, EKF can not estimate a non-differntiable model. I only estimate the SV-ZLB model by PF with a $T = 200$ simulated data. To mimic the shape of interest rates in reality, we set the true values of the parameters in the SV-ZLB model are $\rho_0 = 0$, $\rho_1 = 0.96$, $\phi = 0.9$ and $\sigma^2 = 1$.

Table 1.7 represents the estimation result for the SV-ZLB model via PF. Figure 1.11 illustrates the particle filtered states by fixing the parameters at the value of estimators in table 1.7. $\hat{\rho}_0$ is 0.0539, which is 5% higher than the true value. $\hat{\rho}_1$ is 0.9570, which is 0.3% higher than the true value. $\hat{\phi}$ is 0.9180, which is 2% higher than the true value and $\hat{\sigma}$ is 0.9636, which is 4% lower than the true value. PF estimators for SV-ZLB are unbiased and PF filtered state is close to the true state showed by figure 1.11.

1.5 Conclusion

In this chapter, I make a Monte Carlo study to compare three different simulation based algorithm: Kalman filter, Extended Kalman filter and particle filter. The comparison is via estimating linear Gaussian, non-linear, non-Gaussian, Stochastic Volatility model and Stochastic Volatility model with ZLB. I examine the accuracy of estimated parameter, level of likelihood and bootstrapping variance. I also examine the effect of using different length of data and different number of particles. I found PF is as good as KF when estimating the linear Gaussian model and provide better filtered state than KF for the linear non-Gaussian model. For the SV model, PF provides more accurate estimates for the transition equation than EKF. Finally, PF estimators for the SV-ZLB model are unbiased and filtered state fits true state well.

Chapter 2

On Monetary Policy Uncertainty at the Zero Lower Bound

2.1 Introduction

U.S nominal interest rates have been stuck at zero since 2008. Conventional Gaussian term structure models (GATSM) do not constrain interest rate to be non-negative because that is not necessary while interest rate is far above zero. However, lately in zero interest rate environment, if interest rate models do not assume interest rate to be non-negative but the data are constrained by the zero lower bound, then the models can not validly represent interest rate and its dynamics. Correspondingly, many research of zero interest rates arose. Black (1995) first introduces the concept of shadow rate, which can be a negative value as interest rate is at zero. After that Kim and Singletion (2012), Kripperner (2012), Ichiue and Ueno (2013) and Christensen and Rudebusch (2013) all construct GATSM with a zero lower bound (GATSM-ZLB). GATSM-ZLB corrects the possible negative interest rate in GATSM by introducing shadow rates. Our model follows their way to constrain interest rates to be non-negative. When nominal interest rate is at zero, it has no longer movement. The advantage of including shadow rate in model is we can study on shadow rates' movements instead of interest rates' movements.

However, GATSM-ZLB does not allow for time-varying volatility (heteroskedasticity). Before 2008, nominal interest rates were the main monetary instrument and interest rates' volatility measures the stance of monetary policy. Monetary policy is determined by underlying macro-economic factors and so is volatility. When interest rate stuck at zero, it seems less volatile. Due to underlying macro-economic factors are even more volatile after 2008 financial crisis, volatility are expected to be higher. But previous work either assumes constant volatility or square-root precess, which implies low interest rate must have low volatility. Thus, interest rates models should involve a time-varying volatility which rely on the underlying macro-economic factors. Then the models would be more consistent with the reality that when interest rate is low, volatility is not necessary to be low.

We are the first to construct a stochastic volatility state space model with a zero lower bound (SV-ZLB) to meet the needs of both non-negative interest rate and time-varying volatility. The SV-ZLB model is based on the conventional stochastic volatility (SV) model and constrains the measurement equation to non-negative. SV models have volatility as an unobserved state variable, where the log of the squared residuals is a first-order autoregression (Harvey at al. 1994). The SV model guarantees interest rate volatility is time-varying by letting the volatility depends on the state variable, not the interest rate level itself. We contribute to construct an interest rate model consistent with the fact that low interest rate do not necessarily have low volatility. Meanwhile, the model has no possibility of negative interest rate.

The difficulty of estimating the SV-ZLB model is that the model is a non-linear and non-Gaussian state space model. The conventional method to solve linear state space model is Kalman filter (KF). Extended Kalman filter (EKF) is based on KF to deal with non-linear models, but the model must be differentiable. Alternatively, another simulation-based method particle filter (PF), which can estimated non-linear, non-differentiable and non-Gaussian models. For EKF, people use Taylor expantion to linearize measurement equation by taking first order derivative with respect to state vairable; thus, the model must be differentiable. For particle filter, people use importance sampling to simulate latent state variables, that procedure does not have limitation on model's form. We use one of PF algorithms: continuous Sampling Importance with Resampling (SIR) to estimate parameters via Maximizing log-likelihood. Then, we fix the estimated parameters to filter latent state variables. But, one shortcoming of implementing MLE within continuous SIR is computational burden.

We provide a measure of monetary policy uncertainty in the presence of zero interest rate. That measure comes form our SV-ZLB model estimated latent state variable, which is the volatility of shadow rate. We show decreased model estimated shadow rates coincide with the periods of "Quantitative Easing" policy. Moreover, we compare the SV-ZLB model to SV model and find without zero lower bound, volatility will be underestimated for the periods of zero interest rates. Finally, we find our monetary policy uncertainty index can be applied to asset pricing. It has some predict power on future stock returns and as a risk factor, it is priced in financial market.

The structure of the present paper is as follows. In section 2, we construct a SV-ZLB interest rate state space model. We explain estimation methods in section 3, in which we provide a continuous SIR algorithm for the SV-ZLB model. We also show the estimation results with comparison of SV-ZLB and SV estimated shadow rates and volatilities. In section 4, we apply our monetary uncertainty index to financial asset pricing. Section 5 is conclusion.

2.2 Model

2.2.1 Shadow Rate Model

We assume an short term interest rate model with a zero lower bound. Interest rates can not be negative because people can choose to hold currency, which has zero interest instead of bonds. Black(1995) first introduces a shadow rate model:

$$r_t = max\{s_t, 0\} \tag{2.1}$$

We follow his way to describe shadow rate s_t is the same as nominal short term interest rate as interest rate is positive, otherwise the shadow rate stay below zero. Conventional term structure models do not constrain interest rate to be nonnegative, because as interest rates are far above zero, model estimated results would not be affected. But, since 2008 U.S short term interest rates have been stuck at zero, we must consider an interest rate model with a zero lower bound.

2.2.2 Stochastic Volatility

Conventionally, Federal Reserve affects money market by adjusting interest rate, so interest rates can describe the stance of monetary policy. Meanwhile, the volatility of interest rates can represent monetary uncertainty. However, interest rates already achieve the zero lower bound and zero interest rates have no longer movement. Instead, the Federal Reserve now affects market through increasing money supply. Therefore, zero interest rates can not describe monetary policy stance and the volatility of zero interest rates can not represent monetary uncertainty.

As interest rates lost its function as monetary instrument, we propose to measure monetary policy uncertainty by the volatility of shadow rates. We assume interest rates follow a stochastic volatility process, thus our model allows for time-varying volatility (heteroskedasticity).

$$s_t = \beta_t + e^{x_t/2} u_t,\ u_t \sim N(0,1) \tag{2.2}$$

$$x_t = \phi_0 + \phi_1 x_{t-1} + w_t,\ w_t \sim N(0,\sigma^2) \tag{2.3}$$

where x_t is a latent state variable and it follows a AR(1) process. s_t is a partially observed variable and has a noise u_t, which is assumed in a standard Gaussian distribution. Equation (3) is the transition equation in a state space model and w_t is a state noise, which also is in Gaussian. Therefore, the state variable x_t determines shadow rates' volatility and x_t evolves over time.

2.2.3 Observed Variables

Similar to the Taylor rule, our model try to stipulate how much the Central Bank should charge the nominal interest rate in response to macroeconomic factors. The macroeconomic factors we employ are unemployment rate and inflation rate.

$$\beta_t = \beta_0 + \beta_1 inf_t + \beta_2 unem_t + \beta_3 D * inf_t + \beta_4 D * unem_t \tag{2.4}$$

where, inf_t is the inflation rate, $unem_t$ is the unemployment rate. To observe the Central bank's different responses to macroeconomic factors before and after 2008 financial crisis, we let D denote a dummy variable with zeros before year 2008 and ones otherwise.

Besides nominal interest rate, inflation rate and unemployment rate, we also include an economic policy uncertainty index as one observed variable in the model. We assume the volatility x_t affects the economic policy uncertainty index v_t as

$$v_t = \alpha_0 + \alpha_1 v_{t-1} + \alpha_2 x_t + e_t, \ e_t \sim N(0, \gamma^2) \tag{2.5}$$

Baker, Bloom and Davis (2011) build the index from components like the frequency of references to economic uncertainty and policy in newspapers, the number of federal tax code provisions and extent of disagreement among economic forecasters. So, the economic policy uncertainty index is wide and x_t particularly referring to the monetary policy uncertainty.

2.2.4 Model Comparison

2.2.4.1 A stochastic volatility model with zero lower bound vs. without zero lower bound

Equation (2.1)-(2.5) together is a stochastic volatility state space model with zero lower bound (SV-ZLB). It describes the policy rule of interest rate as which is being at zero lower bound. A

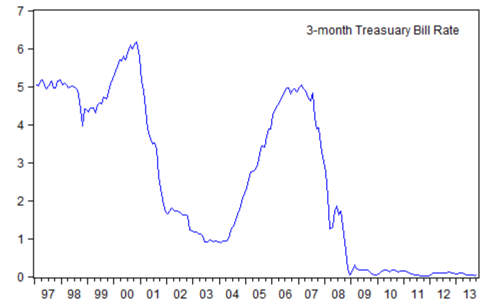


Figure 2.1: 3-month Treasury Bill Rate

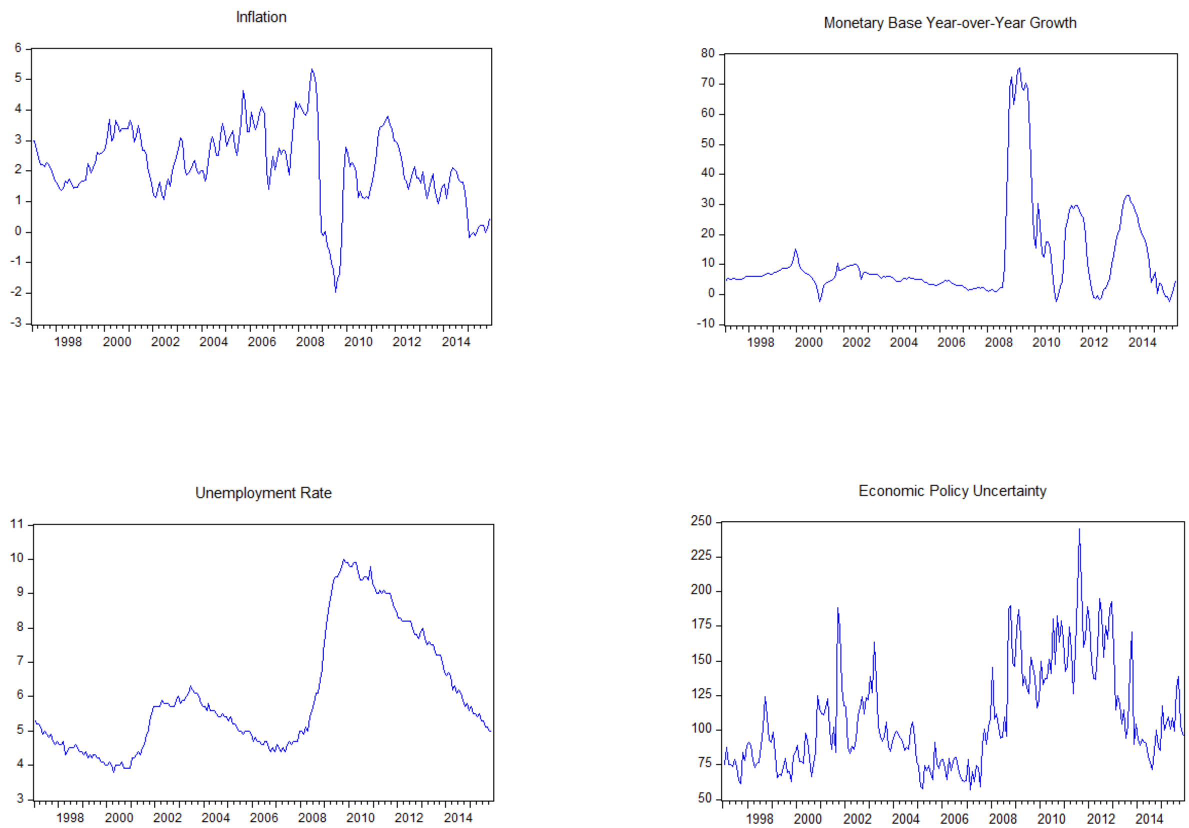


Figure 2.2: Inflation: Year over year growth rate of CPI. Monetary growth: Year over year growth rate of monetary base. Unemployment rate: Percentage of unemployment rate. Economic policy uncertainty: Economic policy uncertainty index from Baker, Bloom and Davis (2001).

conventional stochastic volatility model (SV) has

$$r_t \quad = \quad \beta_t + e^{x_t/2} u_t, \;\; u_t \sim N(0,1) \tag{2.6}$$

instead of equation (2.2). We will compare the estimation result of the SV-ZLB model and the SV model to investigate the consequence of ignoring zero lower bound in zero interest rate environment.

2.2.4.2 A stochastic volatility model with money supply vs. without money supply

Figure (2.1) and (2.2) shows monetary base growth rate increases dramatically when interest rate is down to zero. As interest rate is at zero and the Federal Reserve started to use money supply as the main instrument, the monetary policy stance can be reflected by the volatility of money supply. To allow that circumstances, we extend the interest model with money supply shown as the following equations.

$$
\begin{aligned}
r_t &= max\{s_t, 0\} \\
s_t &= \beta_0 + \beta_1 inf_t + \beta_2 unem_t + \beta_3 D * inf_t + \beta_4 D * unem_t + e^{x_t/2} w_t \;\; w_t \sim N(0,1) \\
V_t &= \alpha_0 + \alpha_1 V_{t-1} + \alpha_2 x_t + e_t \;\; e_t \sim N(0, \lambda^2) \\
M_t &= \gamma_0 + \gamma_1 inf_t + \gamma_2 unem_t + \gamma_3 ip_t + \gamma_4 TDinf_t + \gamma_5 TDunem_t + \gamma_6 TDip_t \\
&\quad + \gamma_7 e^{x_t/2} w_t + \eta_t \;\; \eta_t \sim N(0, \varepsilon^2) \\
x_t &= \phi_0 + \phi_1 x_{t-1} + u_t \;\; u_t \sim N(0, \sigma^2)
\end{aligned}
$$

M_t represents monetary base growth rate. Both shadow rate and money supply are in response to the same macro factors. And the money policy uncertainty is extracted jointly from nominal interest rate and money supply.

2.3 Estimation

2.3.1 Estimation method

2.3.1.1 Filtering latent state variables

Equation (2.1)-(2.5) is the SV-ZLB model we want to estimate. The SV-ZLB model is non-linear, non-differentiable and non-Gaussian, so we implement particle filter to the model. Particle filter is a Monte Carlo simulation-based method for filtering latent state variables by recursively drawing N particles to approximate states variables at each of the time t (see Doucet, Freitas and Gordon 2001). Since, the posterior density $p(x_{0:t}|y_{1:t})$ is unknown where y_t denotes all observations at t, thus which is called target distribution, particles are drawn from a proposal distribution $g_{0:t}(x_{0:t}|y_{1:t};\phi)$ to mimic target distribution, where ϕ denotes all the parameters in the model. To correct the discrepancy between target distribution and proposal distribution, we resample particles from their weights w_t.

The particle filter algorithm we use in this paper is sampling importance resampling (SIR). The detailed SIR procedure for the SV-ZLB model is in Appendix. For SIR algorithm, at time t with given observations and particles at $t-1$, the proposal distribution is $p(x_t|x_{t-1})$, from which we draw particles x_t^i. SIR algorithm assumes $p(x_t|x_{t-1}) = g_t(x_t|x_{t-1}, r_{1:t};\phi)$ and by applying Baye's rule, the weight of particle x_t^i is equal to w_{t-1}^i multiplied by an incremental weight $\tilde{w}_t^i = p(y_t|x_t;\phi)$. So, both proposal distribution and incremental weight are assumed in the state space model. Finally, after resampling particles from their weights, each particle has the same weight $1/N$.

Degenerate problem If without resampling step, calculating particle's weight shows particles accumulate weights over time, so some particles with big weight would grow bigger and some particles with small weight would grow smaller. All the probability mass eventually would be allocated to one particle as the number of iterations increases, that phenomenon is called the degenerate problem (Chopin 2004). With the resampling step, a smaller weight particle will become a smaller numbers of particles and a bigger weight particle will become a bigger numbers of particles, but all of the particles have the equalized weight. So, weights no longer accumulate but convert to numbers of particles. Thus, the SIR algorithm can help to alleviate the degenerate problem from not losing some of the particles.

2.3.1.2 Estimating parameters

We estimate parameters by maximizing log-likelihood (MLE). SIR recursive steps provides log-likelihood at each time t as the mean of weights before resampling. We approximate the continuous log-likelihood by a discrete distribution of particles. Intuitively, observations' likelihood is that conditional on the particle we drew how likely we see the observations. However, the problem of using MLE is that SIR resampling step causes discontinuities in log-likelihood with respect to parameter. Malik and Pitt (2011) provide a continuous resampling procedure to smooth the log-likelihood. We followed their way of smoothing for using MLE to estimate parameters. Once, we obtain estimated parameters, then fixed parameters and run SIR algorithm again to filter latent state variables.

2.3.2 Data and estimation result

We focus on the monthly data of U.S 3-month Treasury Bill rate from January, 1997 through October, 2013. Inflation rate is calculated by taking yearly percentage change of monthly CPI index. Along with unemployment rate, all data are downloaded from the web site of Federal Reserve Bank of St. Louis. We have a total of $T = 202$ observations.

Table 2.1 shows parameter estimates for the SV-ZLB model by using 509 times of continuous SIR. Seeds of drawing random variables are changed for each time of running SIR. Estimators in table 2.1 are the mean of 509 estimates for equation (2.1)-(2.5). Standard error is taking standard deviation of 509 parameter estimates divided by $\sqrt{509}$. Figure 2.3 shows 95% confidence interval of estimated state variable x_t. Series of estimated state variable x_t are obtained by running SIR

	SV-ZLB		SV	
	mean	std deviation	mean	std deviation
β_0	13.3237	0.1350	12.7299	0.1653
$\beta_1\,(\pi_t)$	0.1580	0.0320	0.2843	0.0653
$\beta_2\,(U_t)$	-2.0629	0.0233	-1.9706	0.0255
$\beta_3\,(D\pi_t)$	-1.9243	0.0562	-0.8940	0.1225
$\beta_4\,(DU_t)$	0.5533	0.0115	0.6487	0.0171
ϕ_0	0.0283	0.0253	0.1669	0.0713
$\phi_1\,(x_{t-1})$	0.9779	0.0199	0.9268	0.0179
σ	0.5569	0.0748	0.6144	0.0538
α_0	19.1149	0.6063	19.5525	6.2638
$\alpha_1\,(V_{t-1})$	0.8645	0.0939	1.5419	0.3034
$\alpha_2\,(x_t)$	1.9116	0.5432	1.7333	0.5809
λ	17.0755	0.4117	17.1450	2.0249
LL	-192.44	1.0452	-227.70	1.45

Table 2.1: Estimation result of the SV-ZLB model and SV model without money supply by running 509 times of continuous SIR. Monthly data are from $Jan, 1997$ to $Oct, 2013$. Number of particles is 100.

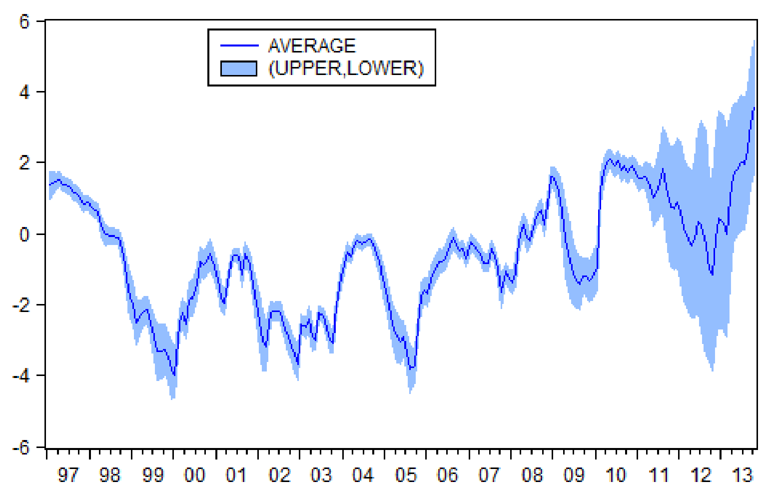


Figure 2.3: 95% confidence interval of estimated state variable x_t by running 509 times of continuous SIR. Monthly data are from $Jan, 1997$ to $Oct, 2013$. Number of particles is 100.

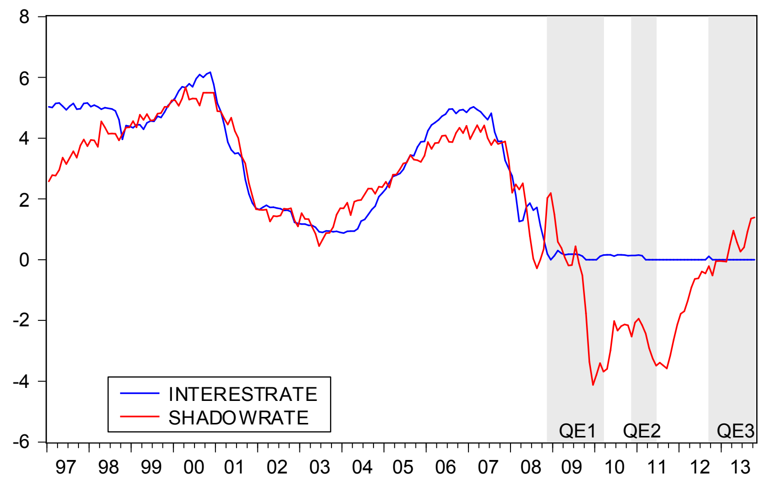


Figure 2.4: SV-ZLB model estimated shadow rate vs. 3-month Treasury Bill rate. The blue line is nominal interest rate and the red line is estimated shadow rate for SV-ZLB model.

with fixing parameters on the parameter estimates.

Table 2.1 indicates shadow rate relate positively to inflation rate and negatively to unemployment rate before year 2008 and the estimates are 0.16 and -2.06 respectively. After 2008, the estimated parameter of inflation rate is -1.76 and the estimated parameter of unemployment rate is -1.5. As inflation or unemployment rate increases, nominal interest rate are reduced after 2008 financial crisis. The effect of inflation is magnified and changed to negative sign could due to endogeneity. After 2008, to boost economy, central bank increases money supply, consequently shadow rate decreases and price level increases.

Figure 2.4 depicts the SV-ZLB model implied shadow rate vs. 3-month Treasury Bill rate. The shaded areas in figure 4 represent the first, second and third "Quantitative Easing" policy period. SV-ZLB implied shadow rate first time began to drop substantially in November 2008 following FOMC's announcement of "QE1". The shadow rate continuously decline in January 2011 corresponding with "QE2". Another dropping of the shadow rate occurs in second half of 2013 when was half year after the announcement of "QE3". So, the SV-ZLB model implied shadow rate shows the unconventional monetary policy "Quantitative Easing" did cause decreasing

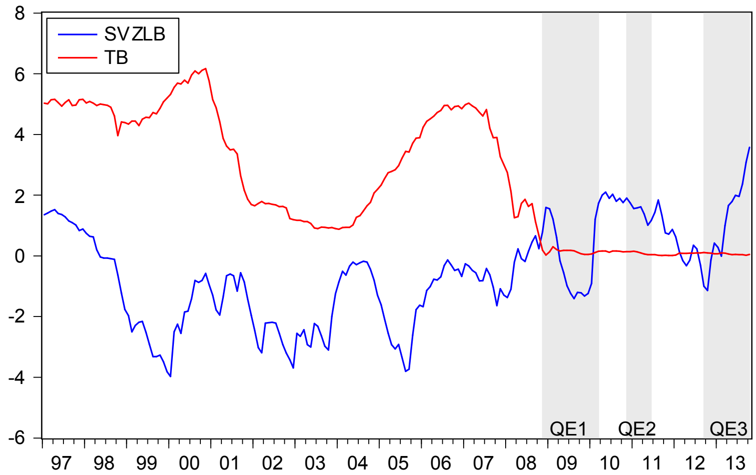


Figure 2.5: SV-ZLB model estimated volatility x_t vs. 3-month Treasury Bill rate. The red line is nominal interest rate and the blue line is estimated volatility of shadow rate for SV-ZLB model.

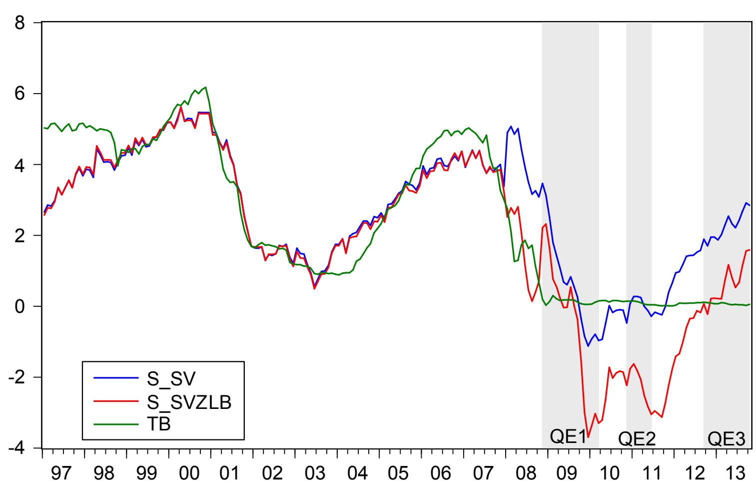


Figure 2.6: Shadow rate for model SV-ZLB vs. interest rate for model SV. The green line is nominal interest rate, blue line is estimated interest rate for SV model and red line is estimated shadow rate for SV-ZLB model.

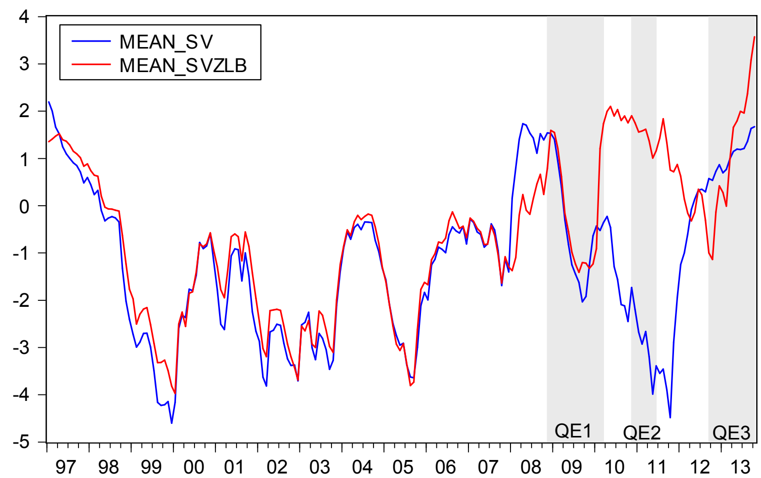


Figure 2.7: Volatility for model SV-ZLB vs. volatility for model SV. The red line is estimated volatility of shadow rate for SV-ZLB model and the blue line is estimated volatility of interest rate for SV model.

in shadow rate. The shadow rate could be a potential indicator for unconventional monetary policy instead of nominal interest rate as it is at zero lower bound.

2.3.2.1 Monetary policy uncertainty

While nominal interest rate is far above zero lower bound, the volatility of interest rate can be used as an indicator of monetary policy uncertainty, because nominal interest rate was the main conventional monetary instrument. For unconventional monetary policy after 2008, nominal interest rate is no longer volatile as it is stuck at zero lower bound, thus we need an new indicator of monetary policy uncertainty. We propose that the volatility of shadow rate in the SV-ZLB model can be served as an new monetary uncertainty index. Figure 2.5 shows the SV-ZLB model estimated latent state variable x_t, which also is the volatility of shadow rate. The advantage of using shadow rate's volatility as monetary policy uncertainty index particularly for unconventional monetary policy is that the movements of shadow rate are not constrained by zero lower bound.

In figure 2.5, the first interesting point is that when nominal interest rate dropped to zero, the volatility did not fall but arose. Unlike what square-root Gaussian term structure models assume is that as nominal interest rate is low, the volatility is low, the fact is in opposite. In 2007, x_t is around 0, while nominal interest rate is 5%. At the end of 2008, x_t is around 2, while nominal interest rate is around 0%. So, conventional Gaussian term structure models can not solve the problem that as zero lower bound is binding nominal interest rate, monetary policy maker should be more uncertain than usual. For this point, the SV-ZLB model does not constrain positive relationship between the level and the volatility of interest rate.

The second interesting point in figure 2.5 is the peaks of volatility coincide with the beginnings and ends of "QE" policies. The volatility is relatively low in the middle of "QE" policy period. At the beginning and end of "QE1", the volatility is around 2 and in the middle, the volatility is around -2. For "QE2", the volatility is around 2 at the beginning, 1 in the middle and 1.7 at the end. That may because monetary policy maker are uncertain about what to do next and when to start or end the policy. As new policy be implemented, policy maker should become less uncertain.

2.3.2.2 Comparison of model SV-ZLB with SV

One question about using the SV-ZLB model to filter monetary policy uncertainty is why can not we ignore the zero lower bound and use the SV model directly. What is the consequence of ignoring zero lower bound for unconventional monetary policy as interest rate is stuck at zero? Theoretically, as Black (1995) introduced because people can choose currency instead of a bond with negative interest rate, interest rate always stay above zero. But conventional Gaussian terms structure models let negative interest rate be possible. Then, our question is empirically what is the difference in estimation result when adding zero lower bound in the model and without.

If we ignore zero lower bound and estimate the SV model, the estimated interest rate does not match with nominal interest rate. Figure 2.6 shows the nominal interest rate, estimated shadow rate for the SV-ZLB model and estimated nominal interest rate for the SV model. Due to the macro-economic variables are still moving when interest rate is stuck at zero, the estimated interest rate for the SV model is volatile and become -1% in the second half of 2009 and soar hight to 3% in

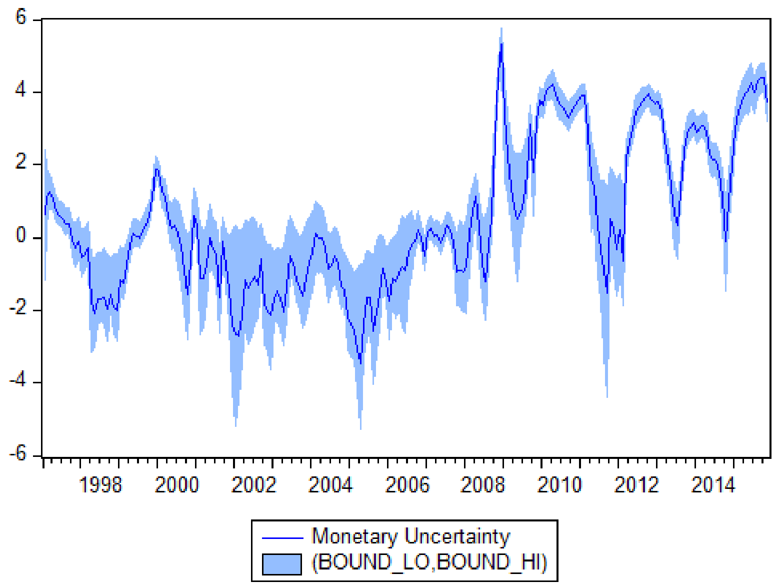


Figure 2.8: SV-ZLB model with money 95 and 5 percentile of Monetary Policy Uncertainty estimated by using 500 bootstrapping estimators.

October 2013. So, the estimated interest rate for the SV model can not fit well with real data after 2008. Comparatively, the SV-ZLB model allows for negative shadow rate and does not constrain shadow rate to move with macro-economic variables.

Moreover, if without zero lower bound, the volatility would be underestimated. Figure 2.7 shows the difference in estimated volatility for the SV-ZLB model and for the SV model. Before 2008, the two volatilities are very close to each other. The biggest difference comes during "QE2" period when SV-ZLB volatility is around 1.5 and the SV volatility is around -3. Without zero lower bound, in the SV model, x_t is the interest rate's volatility, due to that interest rate is less volatile at zero, the estimated volatility is lower than the estimated volatility of shadow rate for the SV-ZLB model. Thus, for unconventional monetary policy, the SV model estimated volatility can not accurately represent the uncertainty.

	SV-ZLB		SV	
	Mean	Std Error	Mean	Std Error
β_0	15.08	0.0403	14.05	0.0362
β_1	-0.03	0.0105	0.00	0.0088
β_2	-2.34	0.0098	-2.17	0.0078
β_3	-2.70	0.0251	-0.02	0.0174
β_4	0.39	0.0075	0.59	0.0062
ϕ_0	0.09	0.0106	0.07	0.0198
ϕ_1	0.89	0.0400	0.78	0.0098
σ	1.33	0.0201	1.71	0.0261
α_0	20.16	0.0892	20.27	0.1126
α_1	0.80	0.0010	0.82	0.0011
α_2	0.08	0.0281	0.04	0.0297
λ	18.88	0.0386	19.19	0.0425
γ_0	6.35	0.0894	6.54	0.1325
γ_1	-1.53	0.0283	-1.61	0.0384
γ_2	0.68	0.0245	0.68	0.0299
γ_3	-0.13	0.0050	-0.11	0.0052
γ_4	4.30	0.0470	6.26	0.0704
γ_5	1.45	0.0207	0.71	0.0263
γ_6	-2.93	0.0163	-3.45	0.0112
γ_7	2.99	0.0039	4.38	0.0033
ε	-0.05	0.0070	-0.008	0.0026
LL	-1975	10.6	-2186	12.26

Table 2.2: Estimation result of the SV-ZLB model and SV model with money supply by running 522 times bootstrapping with random error terms of x_t, M_t and V_t. Monthly data are from Jan, 1997 to Nov, 2015. Number of particles is 100.

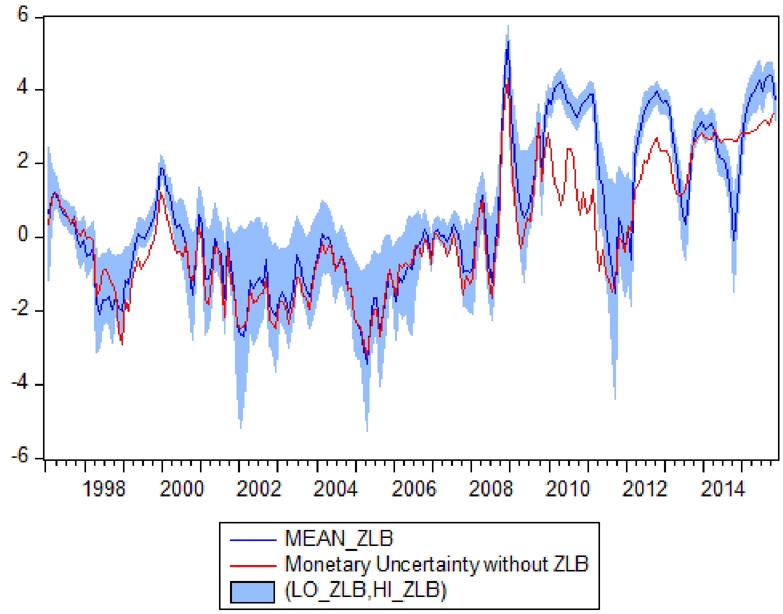


Figure 2.9: The extracted monetary policy uncertainty from the SV-ZLB model with money and the SV model with money

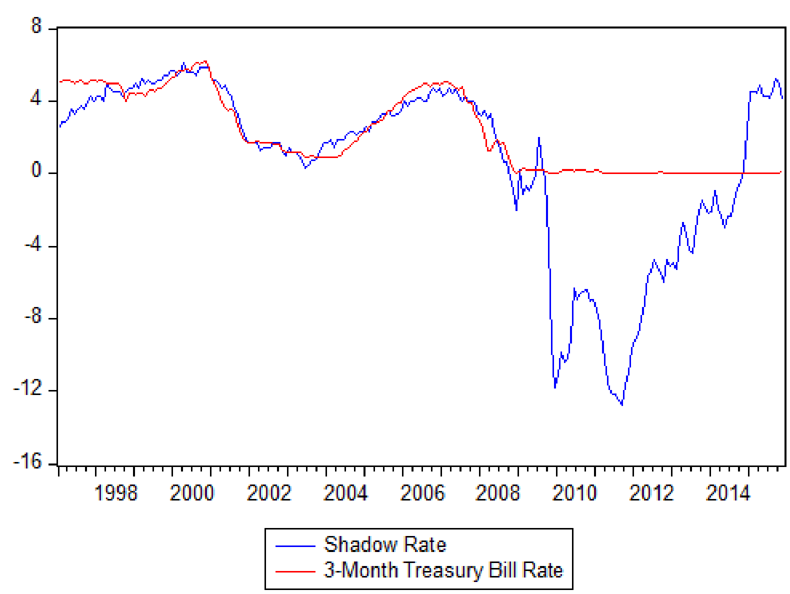


Figure 2.10: SV-ZLB model with money estimated shadow rate and Three-month Treasury Bill rate.

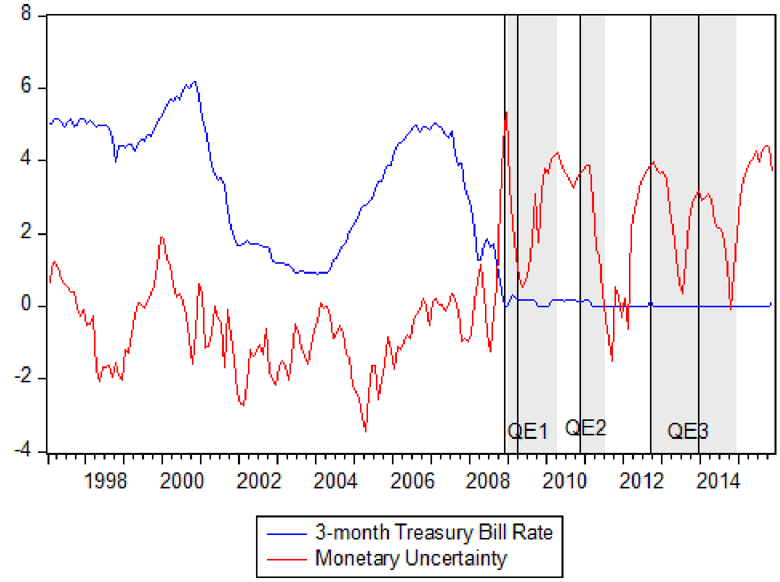


Figure 2.11: Minimum and Maximum of particles filtered by the SV-ZLB model with money using 500 bootstrap estimators.

2.3.2.3 Comparison of model with model without money vs. with money

When the interest rate is at zero, the monetary policy uncertainty is extracted by the SV-ZLB model based on limited information. Those information are macro factors, historical pattern of how shadowrate responding to macro factors, and partial observation that shadow rate is negative. Comparatively, the SV-ZLB model with money has more complete information to extract monetary policy uncertainty. Counter-cyclically, the monetary growth rate moves dramatically when the interest rate is at zero. Therefore, when there is no interest rate movements, the SV-ZLB model with money still can utilize the information of money supply and the historical pattern of how money supply responding to macro factors. Table 2.2 and Figure 2.8-2.11 present the empirical result of estimating the SV-ZLB model with money.

Figure 2.8 shows the monetary policy uncertainty implied by the SV-ZLB model with money. It validates the result of the model without money that the monetary policy uncertainty increases after 2008. Figure 2.11 indicates the uncertainty with Quantitative Easing periods and purchase plan announcement dates. We found that by taking account of money supply, the monetary policy uncertainty rises before each purchase plan begins and trends down after the purchase plan announcement dates. The Quantitative Easing periods and purchase plan announcement dates are listed below.

QE1 (2008Dec-2010Mar)

On November 25, 2008, the Federal Reserve announced that it would purchase up to $600 billion. On March 18, 2009, the FOMC announced that the program would be expanded by an additional $750 billion.

QE2 (2010Nov-2011Jun)

On November 3, 2010, the Fed announced that it would purchase $600 billion of longer dated treasuries, at a rate of $75 billion per month.

QE3 (2012Sep-2014Nov)

On September 13, 2012, the Federal Reserve announced a third round of quantitative easing (QE3). This new round of quantitative easing provided for an open-ended commitment to purchase $40 billion agency mortgage-backed securities per month until the labor market improves "substantially".

The Federal Open Market Committee voted to expand its quantitative easing program further on December 12, 2012. This round continued to authorize up to $40 billion worth of agency mortgage-backed securities per month and added $45 billion worth of longer-term Treasury securities. On December 18, 2013 the Federal Reserve Open Market Committee announced they would be tapering back on QE3 at a rate of $10 billion at each meeting.

The Federal Reserve ended its monthly asset purchases program (QE3) in October 2014.

2.4 Macro Implications

We consider a four variable VAR with 12 lags with the variable ordered as: industrial production growth rate, unemployment rate, Fed Funds rate and the SV-ZLB model implied monetary policy

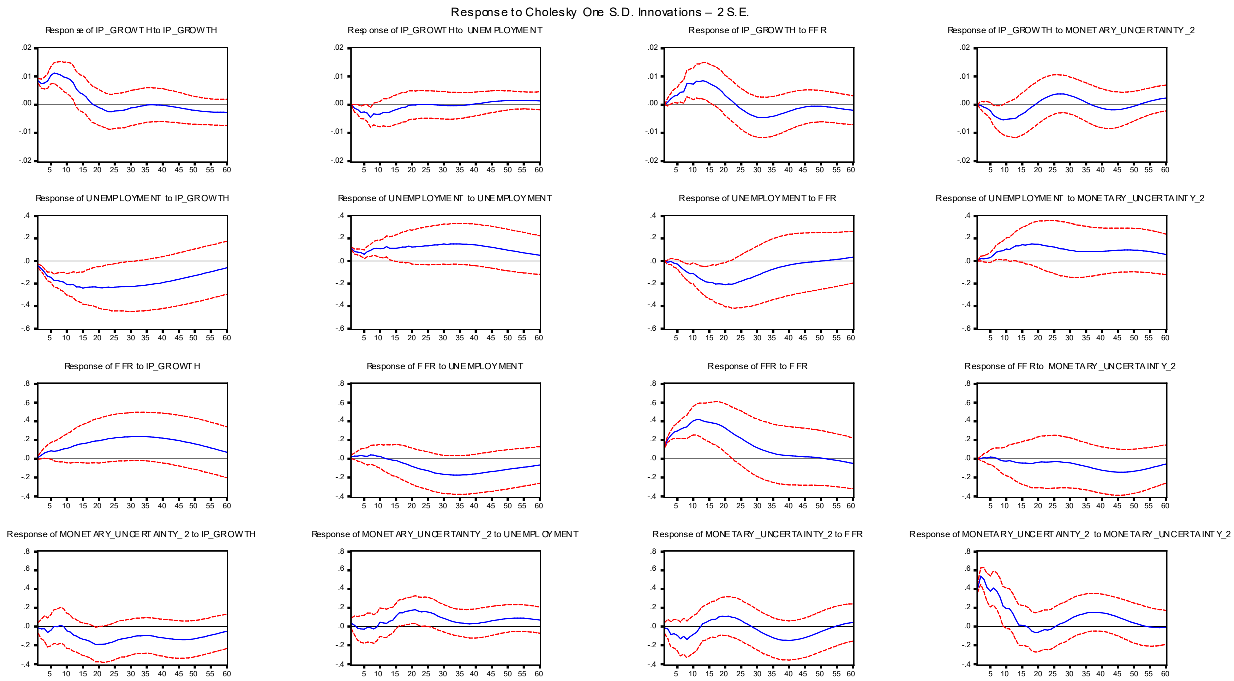


Figure 2.12: Impulse response functions from a four variable VAR(12) with Industrial Production growth rate, Unemployment, Fed Funds rate and Monetary Policy Uncertainty

uncertainty. Impulse responses to one standard deviation shocks are shown in Figure 2.12. The red bands are 2 standard error intervals. We found that one standard deviation shock to the monetary uncertainty lowers the industrial production growth by 0.005% in 10 months and then increases the industrial production growth in 30 months. The uncertainty shock increases the unemployment by 0.15% at its maximum after 15 months. The uncertainty shock reduces Fed Funds rate by 0.14% in 4 years. On the other hand, an output shock lowers the monetary policy uncertainty. An unemployment shock increases the monetary policy uncertainty.

2.5 Asset Pricing Implications

We proposed the volatility x_t in the SV-ZLB model to be a monetary policy uncertainty index and showed how the uncertainty coincide with Quantitative Easing policy. But how can we use the monetary policy uncertainty index? How will it affect asset pricing in financial market? In the following we will prove x_t has predict power for future stock return and x_t is a priced risk factor in market.

2.5.1 Return forecasting

We run the forecasting regression for future S&P500 stock returns on price dividend ratio and volatility by using ordinary least square. The regression is $\sum_{j=1}^{k} \rho^{j-1} r_{t+j} = a + b^{(k)} pd_t + c^{(k)} x_t + \varepsilon_{t+k}$. Let $k = 1, 3, 6, 12$ respectively and $\rho = 0.96$. Excess stock return r_t is calculated by annualized monthly return minus risk free return: $log((D_t + P_t)/P_{t-1}) * 12 - rf_t$. Let price dividend ratio be $pd_t = log(P_t/D_t)$. Monthly stock prices and dividends are from Robert J. Shiller's data base and data are from Jan, 1998 through Oct, 2013. We run these regressions to see whether x_t has predict power on future stock returns.

Table 22 shows forecasting results for 1, 3, 6 and 12 months ahead stock returns on price dividend ratio and volatility. In (b) and (c), x_t is significant for 3 months and 6 months ahead future stock returns. As monetary uncertainty increase by 1, 3 months ahead future return will decrease by 0.09 and 6 months ahead future return will decrease by 0.12. So, monetary policy uncertainty is negatively related to future stock returns. R^2 is bigger if we include more periods stock returns. For $k = 3$, R^2 is 0.1257 and for $k = 6$, R^2 is 0.3598.

2.5.2 Beta pricing models

To see whether volatility is a priced risk factor, we run the Fama-MacBeth two stage regressions. At the first stage, we use 5 years rolling window to run 30-industry portfolio on monetary uncertainty index and on excess market returns respectively: $R_{i,t} - RF_t = \alpha_{xi} + \beta_{xi} x_t + \varepsilon_{xi,t}$ and $R_{i,t} - RF_t = \alpha_{mi} + \beta_{mi}(R_{mt} - RF_t) + \varepsilon_{mi,t}$. For the second stage, we use estimated βs to run regression: $R_i - RF = \alpha + \lambda_m \hat{\beta}_{mi} + \lambda_x \hat{\beta}_{xi} + \eta$. Monthly 30 industries portfolio, risk free rate and market return rates are from Kenneth French data library. Data are from Jan, 1998 through Oct, 2013.

k=1, $R^2 = 0.0409$				
Variable	Coefficient	Std. Error	t-stat	p-value
constant	2.0413	0.7892	2.5865	0.0105
pd_t	-0.4534	0.1932	-2.3471	0.0200
x_t	-0.02035	0.0209	-0.9736	0.3315

(a)

k=3, $R^2 = 0.1257$				
Variable	Coefficient	Std. Error	t-stat	p-value
constant	6.7495	1.3349	5.0561	0.0000
pd_t	-1.5312	0.3242	-4.7237	0.0000
x_t	-0.0865	0.0405	-2.1354	0.0341

(b)

k=6, $R^2 = 0.2223$				
Variable	Coefficient	Std. Error	t-stat	p-value
constant	13.1709	1.6386	8.0377	0.0000
pd_t	-2.9882	0.3929	-7.6048	0.0000
x_t	-0.1198	0.0548	-2.1880	0.0299

(c)

k=12, $R^2 = 0.3598$				
Variable	Coefficient	Std. Error	t-stat	p-value
constant	22.2094	2.0586	10.7887	0.0000
pd_t	-4.9863	0.5130	-9.7196	0.0000
x_t	-0.1091	0.0775	-1.4079	0.1609

(d)

Table 2.3: Forecasting regressions for 1, 3, 6 and 12 periods ahead S&P500 stock returns $\sum_{j=1}^{k} \rho^{j-1} r_{t+j} = a + b^{(k)} pd_t + c^{(k)} x_t + \varepsilon_{t+k}$. We use ordinary least square. $\rho = 0.96$. Annualized excess stock return is $r_t = log((D_t + P_t)/P_{t-1}) * 12 - rf_t$, price dividend ratio is $pd_t = log(P_t/D_t)$ and x_t is our monetary uncertainty induex. Stock prices and dividends are from Robert J. Shiller's website $http://www.econ.yale.edu/\ shiller/data.htm$.

	estimator	std error	t-stat	p-value
$\bar{\alpha}$	0.7523	0.3078	27.5825	0.0000
$\bar{\lambda}_m$	0.1697	0.4826	4.0063	0.0001
$\bar{\lambda}_x$	-0.1120	0.2621	-4.8637	0.0000

Table 2.4: Beta pricing models. First, we use 5 years rolling window to run 30 industries portfolio on monetary uncertainty index and on excess market returns respectively: $R_{i,t} - RF_t = \alpha_{xi} + \beta_{xi} x_t + \varepsilon_{xi,t}$ and $R_{i,t} - RF_t = \alpha_{mi} + \beta_{mi}(R_{mt} - RF_t) + \varepsilon_{mi,t}$. Second, we use estimated βs to run regression: $R_i - RF = \alpha + \lambda_m \hat{\beta}_{mi} + \lambda_x \hat{\beta}_{xi} + \eta$. t-statistic is calculated by the correction method in Shanken(1992). 30 industry portfolio and market returns are from Kenneth French data library $http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html$

In first stage, β_{xi} represents the exposure to monetary policy uncertainty for industry i and β_{mi} represents the exposure to market risk for industry i. Estimated β_{xi} and β_{mi} are two 178×30 matrices. β_{xi} is a negative definite matrix and β_{mi} is a positive definite matrix. That means as monetary policy uncertainty is higher, stock return will be lower and as market excess return is higher, stock return will be higher.

In second stage, for every t, excess returns are regressed on β_{mi} and β_{xi} over 30 industries. Then, estimated price of risk will be averaged over time. $\bar{\lambda}_m$ denotes averaged price of market risk and $\bar{\lambda}_x$ denotes averaged price of monetary policy uncertainty. Table 2.3 shows the result of beta pricing estimation. In table 2.3, both λs are significant. t-statistic is calculated by the correction method in Shanken (1992). The price of market risk is 0.1697 and the price of monetary policy uncertainty is -0.1120. The negative sign is due to estimated exposure to monetary policy uncertainty is negative, so as there is more exposure to monetary policy uncertainty, investors demand extra return to compensate for taking risk. Thus, x_t as a risk factor is priced in financial market.

2.6 Conclusion

We introduce a new interest rate state space model: stochastic volatility model with zero lower bound (SV-ZLB) and provide a particle filter (PF) algorithm: continuous SIR to estimate this kind non-linear and non-Gaussian state space model. We find when interest rate is around zero, shadow rate is below zero lower bound and decreasing trend is correspondent with "Quantitative Easing" policy and the volatility is higher. We also compare SV-ZLB and SV model estimated shadow rate and volatility and find the volatility would be underestimated if we ignore zero lower bound. We suggest to use the estimated volatility of shadow rate as a measure of monetary policy uncertainty. We find it has predict power on 3 and 6 months ahead future stock returns and it is a priced risk factor of portfolios beside price dividend ratio.

Chapter 3

Forecasting Business Cycle Using Yield Curve

3.1 Introduction

The yield curve has become one of the most popular leading indicators of economic activities due to it's cyclical properties. The yield curve is usually upward sloping in the expansion and becomes flattened or inverted before the recession. On one hand, the yield curve contains investors' expectation for the economic status. In recessions, long-term rates intend to be high because investors require higher premium to compensate holding bonds in long term than in short term. And, investors expect short-term interest rate will go up if they are optimistic about future economy. So, long-term rates are higher before expansions and lower before recessions. One the other hand, the short-term rate is the main monetary policy instrument for central bank to adjust economy. Central banks intend to lower the short-term interest rate to stimulate investment in the recession. Therefore, the yield curves and the business cycles have parsimonious relations and should be jointly modeled.

Many literature investigate on the parsimonious relations between yield curves and business cycles. Some researchers model the dynamics of yields jointly with Macro activities through the

Finance-Macro term structure models. Some build probit models of recession in terms of yields curves. Ang, Piazzesi and Wei (2006) construct a VAR of short-term rate, term spread and GDP growth to forecast future GDP growth and estimate market price of risk separately. Their empirical results show that high short rates Granger-cause low GDP growth and that the term spread does not Granger-cause GDP growth. In Stock and Watson (2001), a trivariate model for output growth include lags of output growth, the term spread, and the candidate indicator. They find even though the individual forecasts based on asset prices are unstable, the combined asset price forecast of output growth performs well across the different horizons and countries. Chauvet and Senyuz (2012) use empirical time series proxies of the level, slope and curvature of yield curve from which, they extract a latent yield factor and extract a latent macro factor from monthly industrial production and allow two latent factors to follow a two-state Markov switching processes, which involves cyclical phases of the bond market and the economy. Chauvet and Potter (2001) forecast recessions by estimating a probit model of lagged binary business cycle variable and lagged term spread. All the literature mentioned above try to apply the cyclical properties of yield curve to macro activity forecast. Only Chauvet and Potter (2001) use binary business cycle data instead of ad-hoc macro activity data directly with yield curve to forecast business cycle. However, the probit model in Chauvet and Potter (2001) include the effect of term spread on business cycle but do not consider the effect of business cycle on term spread. Therefore, the previous literature either don't include business cycle data in the term structure models or don't incorporate interrelation between macro variable and yield curve.

I extend the probit model in Chauvet and Potter (2001) to account for two forms of improvement. First, I adopt a yield curve not a term spread to forecast business cycles because previous literature found that the predictive power of the term spread is not stable over time, especially for forecasting the recessions after 1980s. Second, I incorporate the interrelation of business cycles and yields into the model by extracting a latent variable from business cycles and the interrelation of the latent variable and three selected yields. The interrelation allows the yield curve to behave differently according to the economy status and the yield curve to reflect expectation of the future economy. The latent variable is named as the "business cycle factor". This paper uses observed NBER business cycles to represent the overall state of the economy. The advantage of using binary data of business cycles is that it can reduce the dimension of ad-hoc macro variables compared to other macro-finance models.

Furthermore, this paper shows how to estimate the models using Fully Adapted Particle Filtering (FAPF) algorithm. The parameters and latent variables are estimated by maximizing the likelihood, which is generated by FAPF. The likelihood is the product of two parts: the likelihood of observed business cycle and the likelihood of yields. Simulated particles of latent variable are drawn from transition equation and updated by measurement equation. Finally, the likelihood is smoothed by using the method in Malik and Pitt (2011) and maximized by using fmincon function in Matlab. Particle filtering solves the computational difficulties of estimating nonlinear state space model.

The performance of forecasting business cycles is examined under two specifications: the proposed probit model with autocorrelated business cycle factor and a benchmark model without autocorrelated business cycle factor. They are labeled as "Model 1" and "Model 2" respectively. In model 2, the business cycle factor is extracted from observed business cycles and is a portfolio of lagged yields. The distinctive feature of Model 2 is that the effect of lagged business cycle factor on yields is precluded. The empirical results for model 1 shows that an increased business factor by 1 (indicating a higher probability of recession) cause the 3-month rate reduced by 0.106 basis point. And compared to model 2, model 1 has a better in-sample and out-of-sample forecast for the probability of recessions and yields.

The paper is organized as follows. Section 2 explains the data and motivation. Section 3 describes the proposed model with autocorrelated latent business cycle factor and an conditional independent benchmark model. Section 4 identify the parameter restrictions and explains the likelihood maximization estimation through fully adapted particle filter and out-of-sample forecast techniques. Section 5 discuss the empirical results for in-sample and out-of-sample estimations of two models. Section 6 concludes.

3.2 Data and Motivation

This paper measures economic activity by the business cycle data from the National Bureau of Economic Research (NBER). NBER provides the most widely accepted definition of recession: A recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales. A recession begins just after the economy reaches a peak of activity and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion. Therefore, the business cycle is not determined by any single variable but a large collection of variables. Business cycle reflects the change of overall economic activities. By following the previous study, this paper will prove that the yield curve has predicting power for business cycle. In addition, this paper will also emphasize the role of business cycle in forecasting yield curve.

Peak	Trough
December 1969	November 1970
November 1973	March 1975
January 1980	July 1980
July 1981	November 1982
July 1990	March 1991
March 2001	November 2001
December 2007	June 2009

Table 3.1: The NBER business cycle dates from 1969 to 2015

Table 1 lists the NBER business cycle dates from 1969 to 2015. Even with given NBER dates of business cycle, it remains ambiguity about the start date and end date of a recession. Some study (Dieblod and Rudebusch 1989) assumes that the recession starts in the month following the peak month and ends on the date of the subsequent trough month. I use the quarterly data spanning 1969:Q3 to 2015:Q1. So, I follow the convention (Estrella and Trubin 2006), in which a recession

	Start and end dates
Recession 1	1970:Q1 to 1970:Q4
Recession 2	1974:Q1 to 1975:Q1
Recession 3	1980:Q2 to 1980:Q3
Recession 4	1981:Q4 to 1982:Q4
Recession 5	1990:Q4 to 1991:Q1
Recession 6	2001:Q2 to 2001:Q4
Recession 7	2008:Q1 to 2009:Q2

Table 3.2: The start and end dates of recessions from 1969 to 2015

starts in the quarter that directly follows the quarter containing the peak month and that it ends in the quarter containing the trough month. U.S has experienced 7 recessions since 1969. Table 2 shows all the recessions' start quarter and end quarter I use in this paper.

I use the data of end-of-quarter Constant Maturity Treasury (CMT) yields from released Fed H.15 over the period from 1969:Q1 to 2015:Q1. The data in the model are for maturities: 3-month, 5-year and 10-year. Figure 1 shows the yield curve in maturities of 3-month, 1-year, 3-year, 5-year and 10-year. The yield in different maturities spreads out during expansions and close up before recessions. Many research, like (Harvey 1991, 1993) comment that recessions are often preceded by inverted yield curves. Figure 2 shows that every recession since 1969 has been preceded by an negative 10-year term spread except for the recession in 1990.

3.3 Model

3.3.1 State space model representation

This paper uses a model with yield-curve factors, augmented by including observable NBER business cycles. The model is in discrete time. The data are quarterly, so that one period represents one quarter. Let these factors be in a $K+1$ vector F_t, which has K yield-curve factors X_t and 1

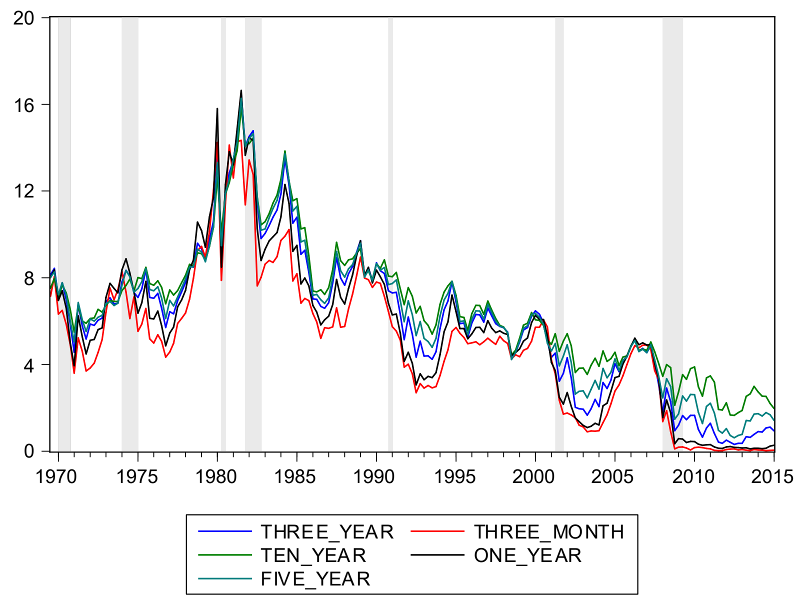


Figure 3.1: Yield curve in maturities of 3-month, 1-year, 3-year, 5-year and 10-year

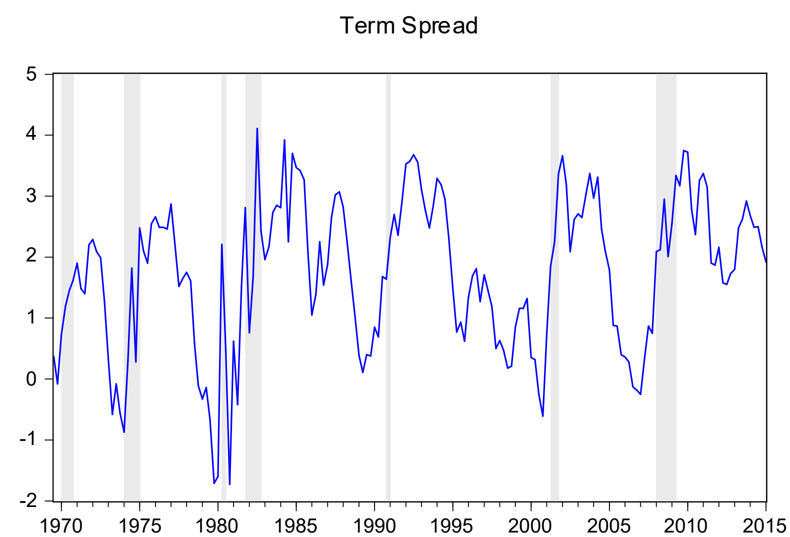


Figure 3.2: The term spread of 10-year bond and 3-month bond

business cycle factor s_t.

$$F_t = \begin{pmatrix} s_t \\ X_t \end{pmatrix} \tag{3.1}$$

As a state space model representation, the vector of state variables follows a Gaussian vector autoregression with lags, showed by a transition equation. The state variables are extracted from observed variables, showed by a measurement equation.

I assume the state variables follow an autoregressive process with a 1-quarter lag and a 4-quarter lag. The assumption is based on the preponderance that the slope of yield curve can predict recessions approximately 4 quarters ahead. To reduce the dimension of parameters, the coefficients of 2-quarter lag and 3-quarter lag of state variables are restricted to zero. The transition equation takes the following form:

$$F_t = \Phi_0 + \Phi_1 F_{t-1} + \Phi_4 F_{t-4} + U_t \tag{3.2}$$

with $U_t \sim$ IID $N(0, \Omega)$. Φ_0 is a $(K+1) \times 1$ vector and Φ_1 and Φ_4 are $(K+1) \times (K+1)$ matrix. Φ_1 and Φ_4 capture the lead-lag relationship between the business cycle factor and yields.

The most efficient way to estimate the model is to assume that all state variables are latent and to use a one-step maximum likelihood estimation. However, this procedure is computationally intensive. I adopt three observed yields in X_t: 3-month, 5-year and 10-year treasury bond interest rates. s_t is the unobserved state variable extracted from business cycles s_t^*. The relation of s_t^* and s_t is described in a probit model:

$$s_t^* = \begin{cases} 1 & if\ s_t > 0 \\ 0 & otherwise \end{cases} \tag{3.3}$$

where a positive s_t indicates a recession and a non-positive s_t indicates an expansion. Therefore, instead of extracting $K+1$ unobserved state variables, replacing the yield-curve factors by observed interest rates reduces computational burden significantly.

3.3.2 Relation to models in the literature

Macro-Finance term structure models The idea of this model is similar to the series of Macro-Finance term structure models. Both model the yield curve's joint dynamics with macro variables. The interrelation of the yield curve and macro variables is as what Rudebusch and Wu (2004) states: central banks move the short rate in response to fundamental macroeconomic shocks should explain movements in the short end of the yield curve; furthermore, with the enforcement by the no-arbitrage assumption, expected future macroeconomic variation should account for movements farther out on the yield curve as well. To model joint dynamics of the yield curve and macro variables, Ang, Piazzesi and Wei (2006) let the vector of state variable, which consist of short rate, term spread and GDP growth, follow an autoregression with one lag. This model has a similar transition equation. However, comparatively the shortcoming of this model is that I do not enforce the no-arbitrage assumption. Therefore, the consistency between long and short rates is not guaranteed by this model.

Probit models of business cycle forecasting This paper following Chauvet and Potter (2001) builds a probit model of recession. The focus of their paper is the impact of term spread on forecasting future recession. In their probit model, the unobserved business cycle factor is estimated by the lagged term spread and the lagged unobserved business cycle factor.

$$s_t^* = \begin{cases} 1 & if\ s_t > 0 \\ 0 & otherwise \end{cases} \tag{3.4}$$

where a positive s_t indicates a recession and a non-positive s_t indicates an expansion.

$$s_t = \beta_0 + \beta_1 X_{t-k} + \theta s_{t-1} + \sigma_t \varepsilon_t \tag{3.5}$$

with $\varepsilon_t \sim$ IID $N(0,\Omega)$. X_t denotes the spread between the 10-year and 3-month Treasury bill rates. K is the forecast horizon of the latent variable in months. The main difference is that their model does not include the impact of business cycle on the term spread. Therefore, in addition to their model, I am able to forecast interest rates as well.

3.4 Estimation

3.4.1 Identification

$$\begin{bmatrix} s_t \\ X_t \end{bmatrix} = \begin{bmatrix} \phi_0^s \\ \phi_0^X \end{bmatrix} + \begin{bmatrix} \phi_1^{ss} & \phi_1^{sX} \\ \phi_1^{Xs} & \phi_1^{XX} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \phi_4^{ss} & \phi_4^{sX} \\ \phi_4^{Xs} & \phi_4^{XX} \end{bmatrix} \begin{bmatrix} s_{t-4} \\ X_{t-4} \end{bmatrix} + \begin{bmatrix} \Sigma^{ss} & \Sigma^{sX} \\ \Sigma^{Xs} & \Sigma^{X} \end{bmatrix} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^X \end{bmatrix} \tag{3.6}$$

$$\varepsilon_t^s \sim N(0,1), \;\; \varepsilon_t^X \sim N(0, I_{k\times 1}) \tag{3.7}$$

Equation (3.6) is the transition equation of the state variable. Φ_1^{ss} and Σ^s are scalars. Φ_0^X and Φ_1^{Xs} are 3×1 vectors. Φ_1^{XX} and Σ^X are 3×3 matrices. I assume in the yield-business cycle model, yield-curve factors only have impact on business cycle in 4-quarter ahead. The business cycle factor only have impact on yield curves in 1-quarter ahead. The error terms of the business cycle factor and yields are uncorrelated. The restrictions are (i)$\phi_1^{sX} = 0$; (ii)$\phi_4^{ss} = 0$; (iii)$\phi_4^{Xs} = 0$; (iv)$\phi_4^{XX} = 0$; (v)$\Sigma^{sX} = 0$; (vi)Σ^{Xs}; (vii)Φ_0^s; (viii) Σ^X is a lower triangular matrix.

The yield-business cycle model will be compared to a benchmark model without autocorrelated business cycle factor. The restrictions are (i) $\phi_1^{ss} = 0$; (ii) $\phi_1^{Xs} = 0$; (iii)$\phi_1^{sX} = 0$; (iv)$\phi_4^{ss} = 0$; (v)$\phi_4^{Xs} = 0$; (vi)$\phi_4^{XX} = 0$; (vii)$\Sigma^{sX} = 0$; (viii)Σ^{Xs}; (ix)Φ_0^s; (x) Σ^X is a lower triangular matrix. The purpose of comparing these two models is to find out whether business cycle and yields are interrelated or only yield curve has predict power for business cycle. Two models will be compared by in-sample and out-of-sample performance of forecasting business cycle and yields.

Yield-Business cycle model restrictions

$$\begin{bmatrix} s_t \\ X_t \end{bmatrix} = \begin{bmatrix} \phi_0^s \\ \phi_0^X \end{bmatrix} + \begin{bmatrix} \phi_1^{ss} & 0 \\ \phi_1^{Xs} & \phi_1^{XX} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & \phi_4^{sX} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s_{t-4} \\ X_{t-4} \end{bmatrix} + \begin{bmatrix} \Sigma^s & 0 \\ 0 & \Sigma^X \end{bmatrix} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^X \end{bmatrix} \tag{3.8}$$

Benchmark model restrictions

$$\begin{bmatrix} s_t \\ X_t \end{bmatrix} = \begin{bmatrix} \phi_0^s \\ \phi_0^X \end{bmatrix} + \begin{bmatrix} 0 \\ \phi_1^{XX} \end{bmatrix} X_{t-1} + \begin{bmatrix} \phi_4^{sX} \\ 0 \end{bmatrix} X_{t-4} + \begin{bmatrix} \Sigma^s & 0 \\ 0 & \Sigma^X \end{bmatrix} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^X \end{bmatrix} \tag{3.9}$$

3.4.2 Fully adapted filtering

Fully adapted particle filtering (FAPF) is adopted to estimate the Yield-Business cycle model. To evaluate the probability of recession, I followed the method in Malik and Pitt (2011).

$$p(s_t^*|s_{t-1}) = \int p(s_t^*|s_t)p(s_t|s_{t-1})ds_t \tag{3.10}$$

$$p(s_t|s_{t-1};s_t^*) = \frac{p(s_t^*|s_t)p(s_t|s_{t-1})}{p(s_t^*|s_{t-1})} \tag{3.11}$$

Equation (3.9) and (3.10) can be viewed as changing the data generating process, in that a sample of the data $s = (s_1, ..., s_T)$ can be obtained. The process involves two steps: sample the latent factor by it's transition equation and resample the latent factor by given next period's information. The detailed algorithm of FAPA is described as following:

Algorithm of FAPF For $t = 0, ..., T-1$:

Given the samples $s_{t-1}^k \sim p(s_{t-1}|I_{t-1})$ for $k = 1, ..., N$. where, I_{t-1} represents all the information at $t-1$.

1. For $i = 1 : N$, compuate

$w_{t-1|t}^i = p(s_t^*|s_{t-1}^k)$ and normalize the weight $\pi_{t-1|t}^i = \frac{w_{t-1|t}^i}{\sum_{i=1}^N w_{t-1|t}^k}$.

2. For $k = 1 : N$, resample $\tilde{s}_{t-1}^k$ from the weight $\pi_{t-1|t}^i$.

3. For $k = 1:N$, sample $s_t^k \sim p(s_t|\tilde{s}_{t-1}^k; s_t^*)$.

3.4.3 Likelihood maximization

The joint likelihood is the product of two parts: the likelihood of observed business cycle and the likelihood of observed yields as equation (10). The likelihood of observed business cycle is generated in the filtering process for business factor in equation (8). The likelihood of observed yields can be calculated by the multivariate normal probability density function, which describes the transition process of cp_t in terms of s_{t-1} and cp_{t-1}. The maximization process is conducted by matlab function fmincon. The initial value of parameters are roughly estimated by running VAR of $[s_t^* \ cp_t]$ on its 1- and 4-quarter lags. By given the initial value of parameters, estimating parameters and filtering latent state s_t can be done at the same time via likelihood maximization.

$$f(y_t|y_{t-1}) = f(s_t^*|s_{t-1}, cp_{t-1}) \times f(cp_t|s_{t-1}, cp_{t-1}) \tag{3.12}$$

$$f(s_t^*|s_{t-1}, cp_{t-1}) = \Phi(\frac{\phi_0^s + \phi_1^{ss} s_t + \phi_4^{sX} X_{t-4}}{\Sigma^s})^{s_t^*} \times [1 - \Phi(\frac{\phi_0^s + \phi_1^{ss} s_t + \phi_4^{sX} X_{t-4}}{\Sigma^s})^{s_t^*}]^{(1-s_t^*)} \tag{3.13}$$

where, Φ is the Cumulative of Distribution Function of the standard normal distribution.

$$f(cp_t|s_{t-1}, cp_{t-1}) = MVN(X_t, \phi_0^X + \phi_1^{Xs} s_{t-1} \phi_1^{XX} X_{t-1}, \Sigma_X \Sigma_X') \tag{3.14}$$

where, MVN represents the multivariate normal probability density function.

3.5 Empirical Estimation

Section 5.1 interprets the parameter estimates of the the model with business cycle factor (model 1) and of the model without business cycle factor (model 2). Section 5.2 explains the method of forecasting in-sample probability of recession and interprets the in-sample forecast performance. Section 5.3 explains the method of forecasting out-of-sample probability of recession and interprets the out-of-sample forecast performance. To examine the effect of adding business cycle factor in the model, results are displayed for both models in each of the section.

3.5.1 Parameter estimates

The model estimates are obtained through maximizing likelihood, which is generated by Fully Adapted particle filtering with 100 particles. Parameters are estimated by using the yield curve data from 1969:Q1 to 2015:Q1 and the NBER business cycle indicator from 1969:Q1 to 2009:Q4. Because the NBER business cycle data are only available until 2009:Q4, the last 21 observations of business cycle are pretended to be zero. Likelihood maximization starts from evaluating the likelihood function at fixed initial value of parameters. To make the initial value of parameters more reasonable, I use the estimates from unconstrained VAR. Then the likelihood function will be maximized by numerically searching different value of parameters.

Table 3.3 and table 3.4 summarizes the parameter estimates of model 1 and model 2 respectively. In table 3.3, the coefficient of one-quarter lag of business cycle factor on the current one is 0.800 and on the 3-month interest rate is -0.106, both are significant. The on-quarter lag of business cycle does not have significant effect on 5-year and 10-year interest rates. The estimates of 4-quarter lag of 3-month interest rate on business cycle is 0.495, which is significant. The 5-year and 10-year interest rates don't have significant forecast power for latent business cycle factor. Surprisingly, the short term interest rate has more predicting power for future business cycle than the interest rate slope. Compared to table 3.3, table 3.4 shows none of the 4-quarter lag of interest rates has effect on the latent business cycle factor. When the effect of latent business cycle factor on interest rates are excluded, the predict power of interest rates on business cycle disappears. All of the standard error of model 1's estimates are smaller than the standard error of model 2's estimates. Besides, Model 1 generates a larger log likelihood than model 2.

	Φ_0	Φ_1				Φ_4			
s_t	-0.046	0.800	-	-	-	-	0.495	-0.370	-0.095
	(0.142)	(0.203)	-	-	-	-	(0.070)	(0.240)	(0.182)
$3-month$	-0.459	-0.106	0.943	-0.145	0.200	-	-	-	-
	(0.333)	(0.042)	(0.095)	(0.361)	(0.325)	-	-	-	-
$5-year$	-0.182	-0.029	0.292	0.033	0.727	-	-	-	-
	(0.262)	(0.039)	(0.076)	(0.286)	(0.256)	-	-	-	-
$10-year$	-0.081	-0.012	0.251	-0.67	1.449	-	-	-	-
	(0.222)	(0.038)	(0.066)	(0.243)	(0.217)	-	-	-	-
	Σ								
s_t	0.896	-	-	-					
	(0.259)								
$3-month$	-	0.967	-	-					
	-	(0.051)							
$5-year$	-	0.593	0.505	-					
	-	(0.049)	(0.028)	-					
$10-year$	-	0.447	0.469	0.148					
	-	(0.044)	(0.028)	(0.011)					
LogL	-326.24								

Table 3.3: Maximum likelihood estimates for model 1

	Φ_0	Φ_1			Φ_4		
s_t	-0.824	-	-	-	0.692	-0.520	-0.092
	(0.540)	-	-	-	(0.374)	(0.592)	(0.500)
$3-month$	-0.139	0.954	-0.223	0.260	-	-	-
	(0.373)	(0.077)	(0.370)	(0.382)	-	-	-
$5-year$	-0.236	0.296	0.006	0.749	-	-	-
	(0.461)	(0.063)	(0.261)	(0.289)	-	-	-
$10-year$	-0.050	0.253	-0.684	1.460	-	-	-
	(0.394)	(0.057)	(0.189)	(0.203)	-	-	-
	Σ						
s_t	0.887	-	-	-			
	(0.418)						
$3-month$	-	0.983	-	-			
	-	(0.176)					
$5-year$	-	0.590	0.507	-			
	-	(0.066)	(0.106)	-			
$10-year$	-	0.442	0.472	0.147			
	-	(0.068)	(0.090)	(0.013)			
LogL	-345.85						

Table 3.4: Maximum likelihood estimates for model 2

Figure 3.3 and figure 3.4 show the estimated business cycle factor with actual NBER business cycle indicator for model 1 and model 2 respectively. The actual recession occurs when the observed business cycle equals to 1. The estimated business cycle factor is compared to the actual business cycle. A positive estimated business cycle factor indicates a recession. A negative estimated business cycle factor indicates an expansion. In figure 3.3, estimated business cycle factor by model 1 matches all 7 recessions from 1969:Q1 to 2015:Q1. The estimates of business cycle by model 2 missed 3 recessions: 1971, 1991 and 2008. Therefore, in terms of accuracy, model 1 can better estimate the recessions happened in the past.

3.5.2 In sample forecasts

The main purpose of building a probit model of business cycles and the yield curve is to forecast business cycle turning points. The probability of a recession is calculated by using the maximum likelihood estimates for the full sample 1969:Q1-2015Q1. Given the estimates, I define the recession probability forecast as

$$Pr(S_t^* = 1|I_{t-1}) = \Phi[\hat{\phi}_0^s + \hat{\phi}_1^{ss} s_{t-1} + \hat{\phi}_4^{sX} X_{t-4}, \hat{\Sigma}_{11}], \tag{3.15}$$

where $\hat{\phi}_0^s$, $\hat{\phi}_1^{ss}$, $\hat{\phi}_4^{sX}$ and $\hat{\Sigma}_{11}$ are the model estimated parameters. The posterior predictive distribution can be calculated and used to present the probability of recessions. In addition to the estimated parameters, to obtain the recession prediction, one needs to integrate out the unknown latent variable s_{t-1}. For $\{t = 1, 2, ...T\}$, this is numerically integrated at each iteration by particle filtering. Then, s_{t-1} is the average of 100 resampled particles.

Figure 3.5 and figure 3.6 display the in-sample estimated probability of recession for model 1 and model 2 respectively. In figure 3.5, all of the estimated probability of recession are bigger than 50% during the 7 recession periods and the estimated probability is persistent and stable. The estimated probability in figure 3.5 is more volatile. The probability is smaller than 50% during 3

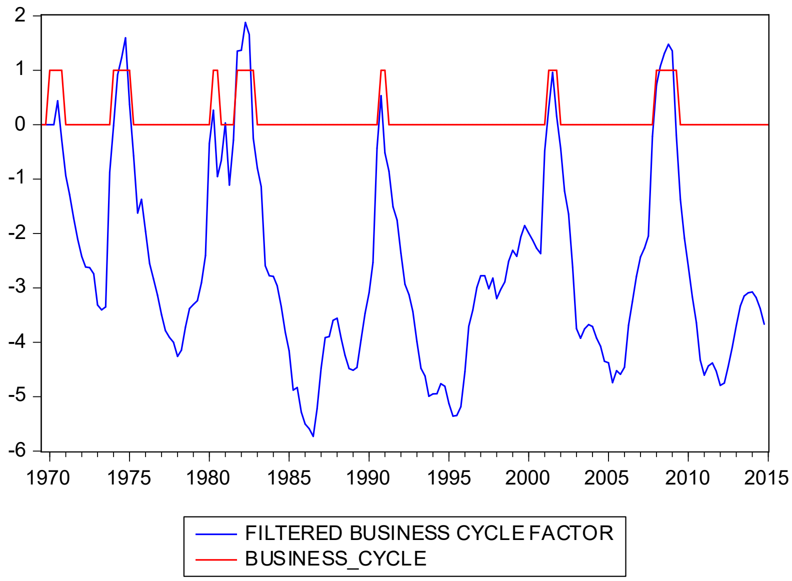


Figure 3.3: Filtered business cycle factor by model 1

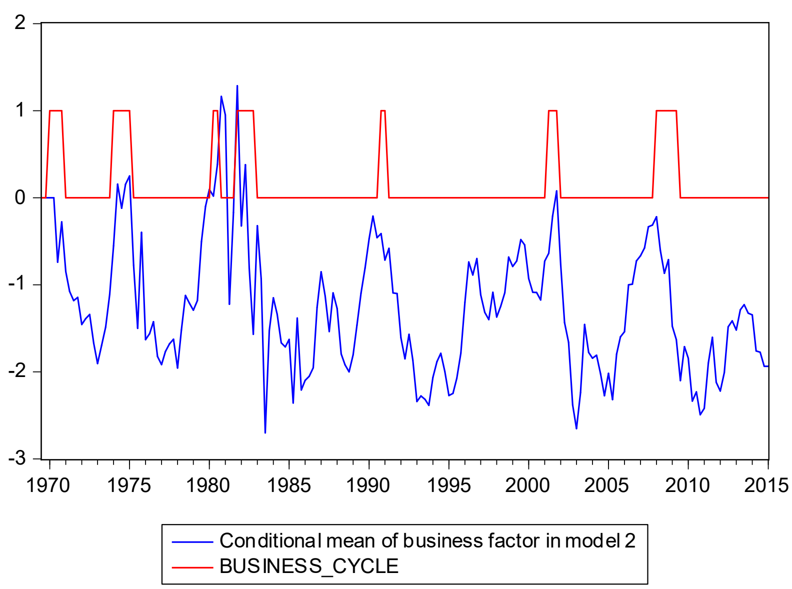


Figure 3.4: Conditional mean of business cycle factor by model 2

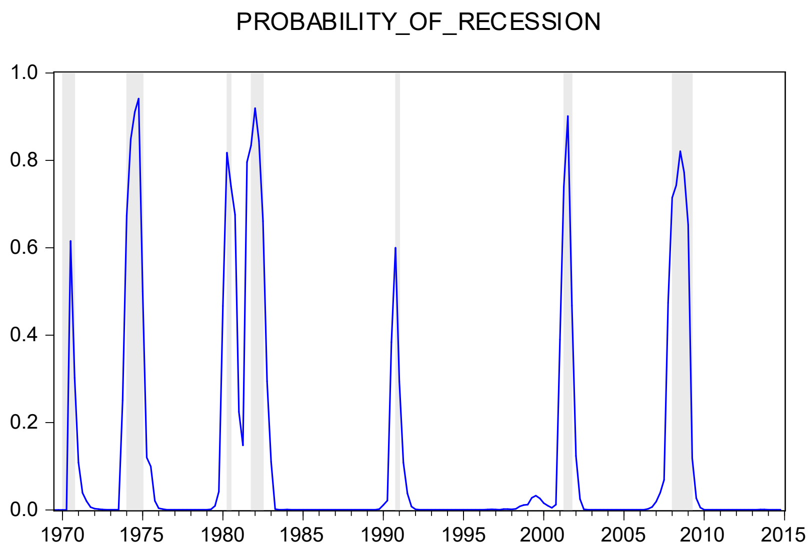


Figure 3.5: In-sample estimated probabilities of recession for model 1

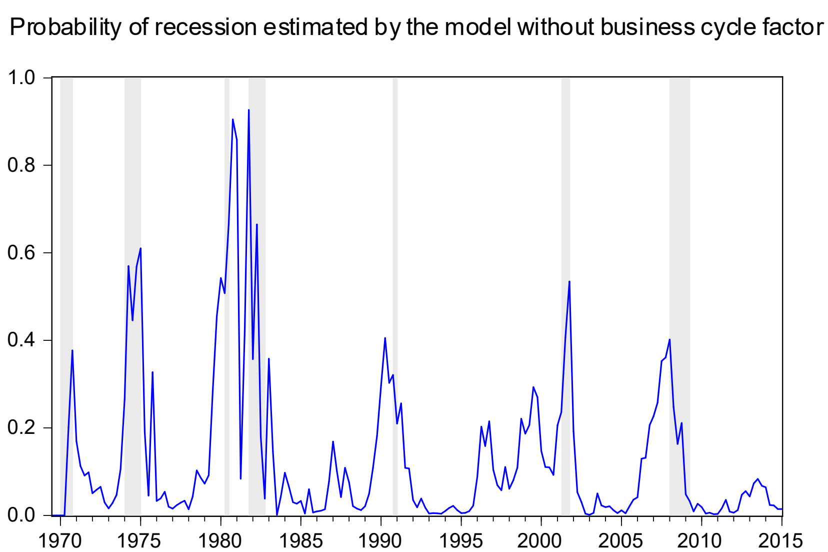


Figure 3.6: In-sample estimated probabilities of recession for model 2

recessions in 1971, 1991 and 2008.

3.5.3 Out-of-sample forecast

The out-of-sample forecasting is to evaluate the performance of forecasting future business cycles and interest rates. I followed the out-of-sample forecast method in Chauvet and Senyuz (2012). The h-step ahead forecasts $p(s^*_{t+h} = 1|I_t)$ uses in-sample estimates and observations, where I_t denotes all information up to t. I aim to forecast for the interval of 2006:Q4 - 2015:Q1, which includes one actual recession in 2008. I consider the forecast horizons from 1 quarter to 4 quarters, $h = 1,2,4$. Two models will be estimated using data from 1969:Q1 to 2006:Q3. The probability of recessions is recursively estimated for each month from 2006:Q4 to 2015:Q1. For model 1, the latent variable s_t will be re-filtered by given information s^*_t for each of the forecasting period.

I evaluate the forecasts using three common measures of forecast accuracy: the mean absolute error (MAE), the log probability score (LPS) and the quadratic probability score (QPS). These measures are evaluated using $h-step$ ahead forecast errors, which is obtained through the recursive forecasting scheme. These measures are defined as:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|p_{t|t-h} - s^*_t| \tag{3.16}$$

$$LPS = \frac{1}{T}\sum_{t=1}^{T}[(1-s^*_t)ln(1-p_{t|t-h}) + s^*_t ln(p_{t|t-h})] \tag{3.17}$$

$$QPS = \frac{2}{T}\sum_{t=1}^{T}(p_{t|t-h} - s^*_t)^2 \tag{3.18}$$

where $p_{t|t-h}$ is the $h-step$ ahead probability of recession estimated using observations available at $t-h$. These measures evaluate the closeness of predicted probabilities from the observed recession realizations. The first measure is standard. The LPS is from the more specialized literature on evaluating probability forecasts (Diebold and Rudebusch 1989). The QPS is the most used probabilistic evaluation test from (Brier and Allen 1951). The measures differ on the relative penalty to

Probability forecast	LPS	MAE	QPS
Current-quarter recession prediction			
Model 1	0.106	0.073	0.066
Model 2	0.288	0.174	0.175
One-quarter-ahead recession prediction			
Model 1	0.394	0.107	0.267
Model 2	0.706	0.193	0.329
Two-quarter-ahead recession prediction			
Model 1	0.516	0.182	0.311
Model 2	0.722	0.194	0.338
Four-quarter-ahead recession prediction			
Model 1	0.642	0.187	0.342
Model 2	0.754	0.191	0.353

Table 3.5: Evaluation of probability forecast

Model	MSE for $\hat{i}_t^1$	MSE for $\hat{i}_t^{20}$	MSE for $\hat{i}_t^{40}$
One-quarter-ahead interest rate prediction			
Model 1	0.1671	0.7557	0.5424
Model 2	0.6307	0.7725	0.5478
Two-quarter-ahead interest rate prediction			
Model 1	1.5122	1.8938	1.2552
Model 2	1.7957	1.9508	1.2759
Four-quarter-ahead interest rate prediction			
Model 1	4.4973	3.9386	2.4406
Model 2	4.9000	4.0279	2.4720

Table 3.6: MSE: out-of-sample forecast for interest rates

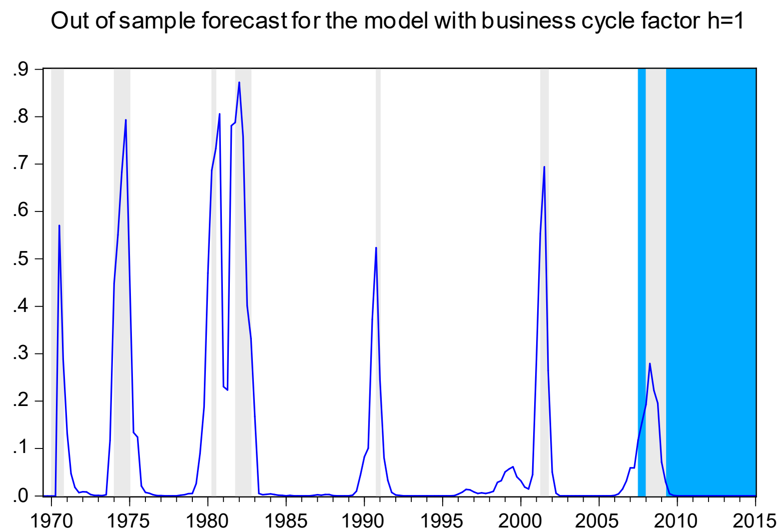


Figure 3.7: Out-of-sample forecast for probabilities of recession by model 1 h=1

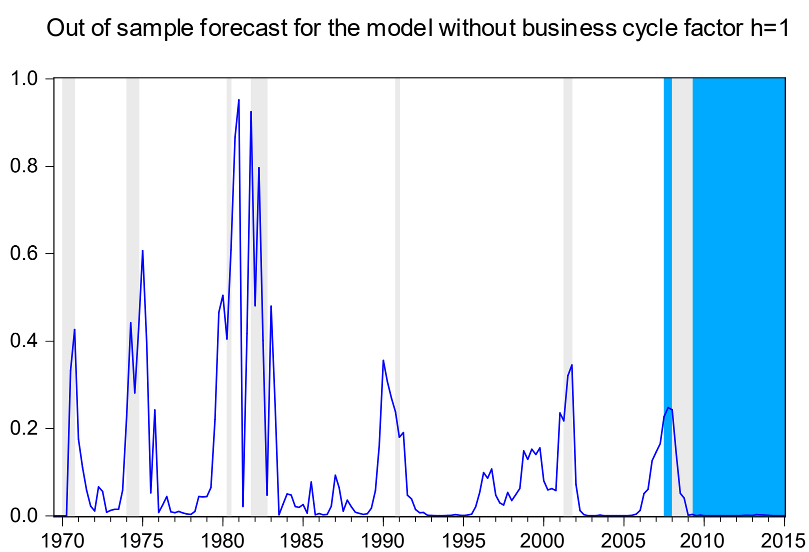


Figure 3.8: Out-of-sample forecast for probabilities of recession by model 2 h=1

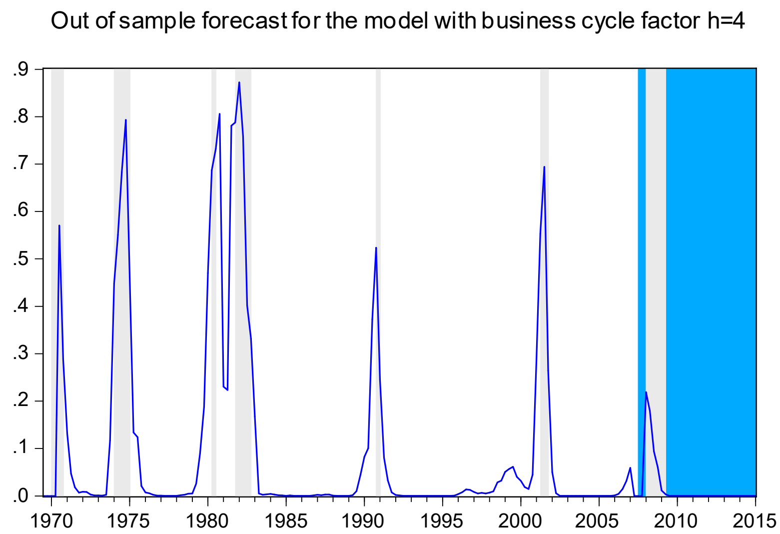


Figure 3.9: Out-of-sample forecast for probabilities of recession by model 1 h=4

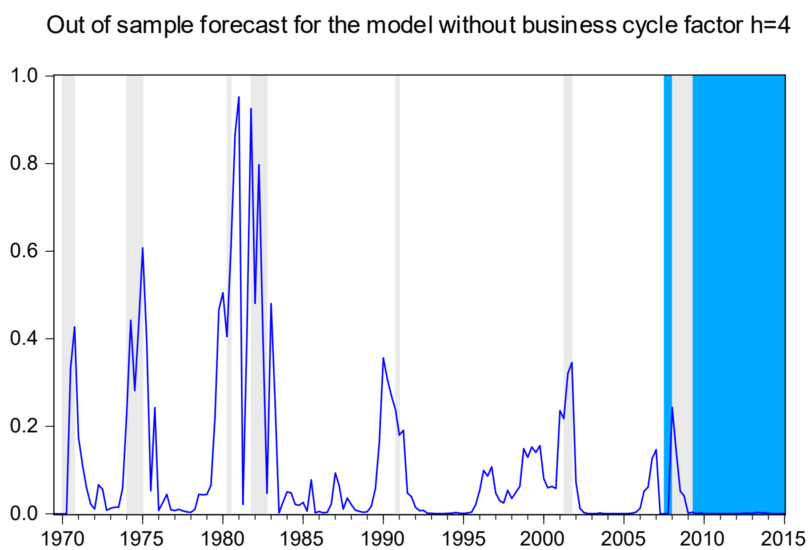


Figure 3.10: Out-of-sample forecast for probabilities of recession by model 2 h=4

large versus small errors. The QPS penalizes larger forecast errors more than smaller ones.

Table 3.5 provides accuracy evaluations for probability of recession. The prediction is built at forecast horizon, $h = 0, 1, 2, 4$. Model 1 has smaller prediction errors at all four forecast horizons. Figure 3.7 to figure 10 present the out-of-sample performance for two models at $h = 1$ and $h = 4$. Forecast periods are shaded in blue and real recessions are shaded in light gray. The estimated probabilities by both model surge at around two years before the 2008 recession and go down to zero when the recession ends. However, the probabilities are not high enough to indicate the 2008 recession. The peak of estimated probability is approximately 30% at $h = 1$ and more than 20% at $h = 4$.

3.6 Conclusions

This paper proposed a probit model that interrelates yield curve and NBER business cycles. Similar to many Macro-Finance models, it emphasizes the interrelationship of interest rates and macro economy. Central bank moves the short rates in response to macroeconomic shocks and long term rate indicates investors' forecast for future economy. This paper's purpose is to forecast business cycles and which is not determined by any single variable but a large collection of economic activities. Therefore, this model directly includes NEBR business cycles as observations instead of uses output measurements to predict business cycles. Meanwhile, I directly use three yields in the model but not only term spread, which is widely used in business cycle forecasting models. This model is explicit as to forecast business cycles and it also provides flexibility for the effect of yields on recession.

In summary, I find (i) short rate has significant effect on business cycle factor 1 year ahead; (ii) the latent business cycle factor has significant effect on forecasting the future short-term interest rate; (iii) long term rate is insignificant on predicting recession. The result shows the proposed model can predict all the turning points of NBER business cycles with no false peaks or troughs and no misssed turns. Moreover, the model with latent business cycle factor and yield curve has better forecasting performance than the model with only yield curve.

References

[1] Ang, A., Piazzesi, M. and Wei, M., 2006, What does the yield curve tell us about GDP growth?, Journal of Econometric, Vol 131, 359-430

[2] Brier, G. W., and R. A. Allen, 1951, Verification of weather forecasts. In Compendium of Meteorology, Malone, T.F., Ed., American Meteorological Society, Boston, pp. 841-848.

[3] Baker, S., Bloom, N. and Davis, S., 2001, Measuring economic policy uncertainty, National Bureau of Economic Research working paper

[4] Chauvet, M. and Potter, S., 2001, Forecasting recessions using the yield curve, Federal Reserve Bank of New York working paper

[5] Chauvet, M. and Zeynep, S., 2012, A dynamic factor model of the yield curve as a predictor of the economy, Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C

[6] Chopin, N., 2004, Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. Annals of Statistics 32, 2358-2411.

[7] Creal, D., 2011, A survey of sequential Monte Carlo methods for economics and finance, Economic Review, pp. 245-96

[8] Christensen, J. and Rudebusch, G., 2013, Estimating Shadow-Rate Term Structure Models with Near-Zero Yields, working paper series, Federal Reserve Bank of San Francisco.

[9] Del Moral, P, 2004, Feynman-Kac Formulae, genealogical and interacting particle systems with applications. Springer, New York.

[10] Diebold, F. and Rudebusch, G., 1989, Scoring the leading indicators, Journal of Business 64(July 1989), 369-391.

[11] Doucet, A., de Freitas, N., Gordon, N.J., 2001. Sequential Monte Carlo methods in practice. Springer-Verlag, New York.

[12] Estrella, A., and M. R. Trubin, 2006, The yield curve as a leading indicator: some practical issues, Federal Reserve Bank of New York, Current Issues in Economics and Finance 12, no.5, July/August.

[13] Filipovic, D. 2004. Term-Structure models a graduate course. Springer Press.

[14] Fischer Black, 1995. Interest rates as option. The Journal of Finance, Vol. 50, No:5, pp. 1371-1376.

[15] Gamerman, D., 1997, Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference, London, UK. SE1 8HN: Champman and Hall.

[16] Harvey, A.C, Ruiz, E. and Shephard, N., 1994, Multivariate Stochastic Variance Models. The Review of Economic Studies, Vol. 61, No. 2, pp. 247-264

[17] Harvey, C.R., 1991, Term structure and world economic growth, Journal of Fixed Income, pp. 7-19

[18] Harvey, C.R., 1993, Term structure forecasts economic growth, Financial Analysts Journal, vol. 49, no.3, pp. 6-8.

[19] Ichiue, H. and Ueno, Y., 2007, Equilibrium interest rate and yield curve in a low interest rate environment, Working paper series, Bank of Japan.

[20] Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering: 35–45.

[21] Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction theory. Journal of Basic Engineering: 95–108.

[22] Kim, C. and Nelson, C., 1999, A Bayesian approach to testing for Markov switching in univariate and dynamic factor models, Discussion papers in Economics at the University of Washington 0035, Department of Economics at the University of Washington.

[23] Kim, D. and Singleton, K., 2012, Term structure models and the zero bound: An Empirical investigation of Japanese yields, Journal of Econometrics,Vol. 170, 32-49.

[24] Krippner,L. 2012, Modifying Gaussian term structure models when interest rates are near the zero lower bound, Discussion Paper, Reserve Bank of New Zealand.

[25] Malik, S., Pitt, M.K., 2011. Particle filters for continuous likelihood evaluation and maximisation. Journal of Econometrics. 165 pp. 190-209.

[26] Mayer, R., Fournier, D.A. and Berg, A., 2003. Stochastic volatility: Bayesian computation using automatic differentiation and extended Kalman filter. Econometrics Journal, Vol, 6, pp. 408-420.

[27] Rubin, D.B., 1988. Using the SIR algorithm to simulate posterior distributions. Bayesian statistics, Vol. 3, pp. 395-402.

[28] Rudebusch, G. and Wu, T., 2004, A macro-finance model of the term structure, Monetary Policy, and the Economy, Federal Reserve Bank of San Francisco working paper 17

[29] Shanken, J., 1992, On the estimation of beta-pricing models, Review of Financial Studies 5, 1-34.

[30] Stock, J.H. and Watson, M.W., 2001, Forecasting output and inflation: the role of asset prices, NBER Working paper 8180

Appendix A

SIR algorithm

A.1 SIR algorithm for the SV model

Given the initial weighted particle set $\{x^i, w^i\}$, for every i:

(i) Simulate $x_t^i \sim \frac{1}{\sigma} exp[-\frac{1}{2\sigma^2}(x_t - \phi x_{t-1}^i)^2]$.

(ii) Compute importance weight: $w_t^i \propto w_{t-1}^i exp[-\frac{(y_t-\rho_0-\rho_1 y_{t-1})^2}{2exp(x_t^i)} - \frac{x_t^i}{2}]$

(iii) Normalize weight: $\hat{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i}$

(iv) Resample N particles with $\{\hat{w}_t^i\}_{i=1}^N$ and set $w_t^i = \frac{1}{N}$.

A.2 SIR algorithm for the SV-ZLB model

Given the initial weighted particle set $\{x_t^i, w_t^i\}$, for every i:

(i) Simulate x_t^i from $\frac{1}{\sigma} exp[-\frac{1}{2\sigma^2}(x_t - \phi x_{t-1})^2]$

(ii)Compute importance weights:

if $y_t > 0 : w_t^i = \frac{1}{\sqrt{2\pi e^{x_t}}} e^{\frac{-(y_t-\rho_0-\rho_1 y_{t-1})^2}{2e^{x_t}}}$

if $y_t = 0 : w_t^i = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi e^{x_t}}} e^{\frac{-(y_t-\rho_0-\rho_1 y_{t-1})^2}{2e^{x_t}}} dy_t$

(iii) Normalize weight: $\hat{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i}$

(iv) Resample N particles with $\{\hat{w}_t^i\}_{i=1}^N$ and set $w_t^i = \frac{1}{N}$.

Appendix B

SIR algorithm for SV-ZLB model

B.1 Filtering latent state variables

Given the initial weighted particle set $\{x_t^i, w_t^i\}$ for $t = 1, ..., t-1$,and $i = 1, ..., N$.

(i) Simulate x_t^i from $p(x_t^i|x_{t-1}^i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x_t^i - \phi_0 - \phi_1 x_{t-1}^i)^2}$

(ii)Compute importance weights:

if $y_t > 0$: $w_t^i = p(r_t|x_t^i) * p(v_t|x_t^i) = \frac{1}{\sqrt{2\pi e^{x_t^i}}} e^{\frac{-(r_t - \beta_t)^2}{2e^{x_t^i}}} * \frac{1}{\sqrt{2\pi\gamma}} e^{\frac{-(v_t - \alpha_0 - \alpha_1 v_{t-1} - \alpha_2 x_t^i)^2}{2\gamma^2}}$

if $y_t = 0$: $w_t^i = p(r_t|x_t^i) * p(v_t|x_t^i) = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi e^{x_t^i}}} e^{\frac{-(r_t - \beta_t)^2}{2e^{x_t^i}}} dr_t * \frac{1}{\sqrt{2\pi\gamma}} e^{\frac{-(v_t - \alpha_0 - \alpha_1 v_{t-1} - \alpha_2 x_t^i)^2}{2\gamma^2}}$

(iii) Normalize weight: $\hat{w}_t^i = \frac{w_t^i}{\sum_{i=1}^N w_t^i}$

(iv) Resample N particles with $\{\hat{w}_t^i\}_{i=1}^N$ and set $w_t^i = \frac{1}{N}$.

B.2 Estimating parameters

Log-likelihood is given by (Malik and Pitt 2001):

$$p(y_t|\theta, y_{1:t-1}) \approx \frac{1}{N} \sum_{i=1}^{N} w_t^i \tag{B.1}$$

$$\mathscr{L}(\theta|y_{1:T})) = \prod_{t=2}^{T} p(y_t|\theta, y_{1:t-1}) = \prod_{t=1}^{T} \frac{1}{N} \sum_{i=1}^{N} w_t^i \tag{B.2}$$

where $y_t = [r_t, v_t]'$.