Can Economic Policy Uncertainty help predict Chinese Stock Market returns? Evidence using an efficient Dynamic Moving Average (eDMA) approach*

Dingqian Liu†

Department of Economics, American University

Abstract

Rational asset pricing theory indicates that stock returns are a function of the real economy. Thus, the fluctuation of the real economy should have a big impact on the stock market, and hence is an essential predictor of stock returns. The stock market of China is sensitive to uncertainty in its economic policy for its policy-driven feature. Using an efficient Dynamic Moving Average (eDMA) model, this paper investigates how well the newspaper-based Economic Policy Uncertainty (EPU) index can predict the returns of the Chinese Shanghai Stock Exchange Index. The empirical evidence shows the impact of monetary policy is muted, when EPU is considered as a predictor. Also, eDMA provides a significant improvement in forecasting performance in comparison to other forecasting methodologies.

JEL classification: C11, C53, G17, Q43

Keywords: Bayesian Methods, Econometric Models, Stock Return Forecasting, Efficient Dynamic Model Averaging, Economic Policy Uncertainty

---

* This is still an extended draft.

†† Corresponding author. Contact information: 4400 Massachusetts Ave., N.W., Washington, DC 20016, USA; Tel: (202) 790-4549; Email: dl5165a@american.edu.

1. Introduction

We are in a world of rising Economic Policy Uncertainty (EPU). The "trade war" between China and the U.S., the rise of populism and uncertainty raised by Brexit, are having big impacts on the global economy. The study of how macro-economic uncertainty influences real economy has risen in the past 10 years. In the seminal work of Bloom (2009), it is shown that under a rising economic uncertainty, companies tend to cut down investment and employment. But how stock market reacts to a rising EPU has not been well investigated. Usually, stock market reacts to news faster than real economy, as the news would affect company's market value, which is quickly priced via trading behavior. When a company decreases investment and employment, it decreases its future expected cash flow and profit. Rational asset pricing theory indicates that real economy fluctuation plays an important role in stock market returns. Literatures have shown a high correlation between EPU and stock market volatility measured by realized volatility or VIX (Bloom 2009, Baker etc. 2016). However, very few research (Brogaard etc. 2015, Pastor etc. 2012) has been made to see if economic uncertainty can help predict stock return. In addition, as shown in Figure 1, the EPU in China shows a rising trend since 2000, which could have a big impact on stock market return.

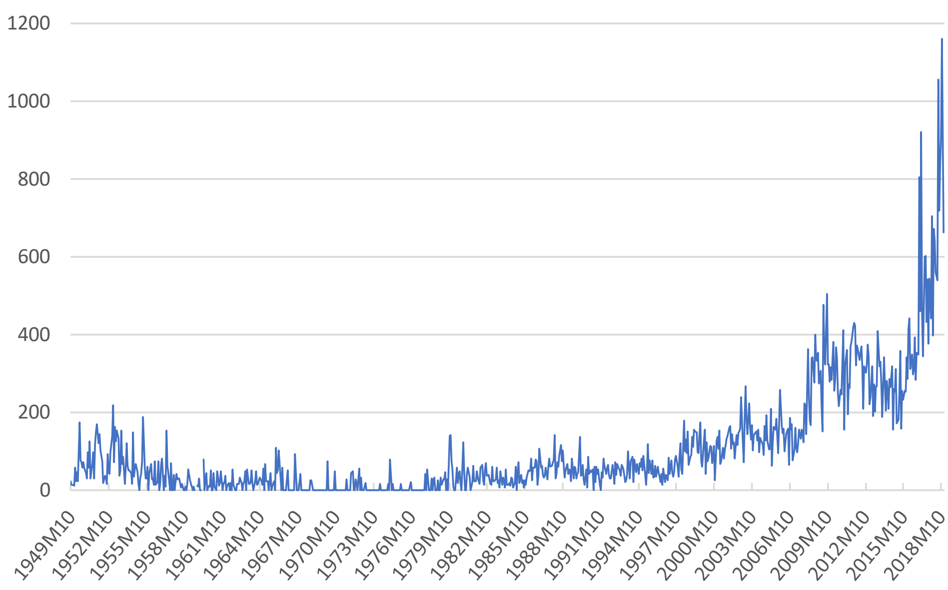

Figure 1 China Monthly EPU Index 1949—2018

Figure 2 shows the main events that drive the spikes of China EPU Index since 2000.

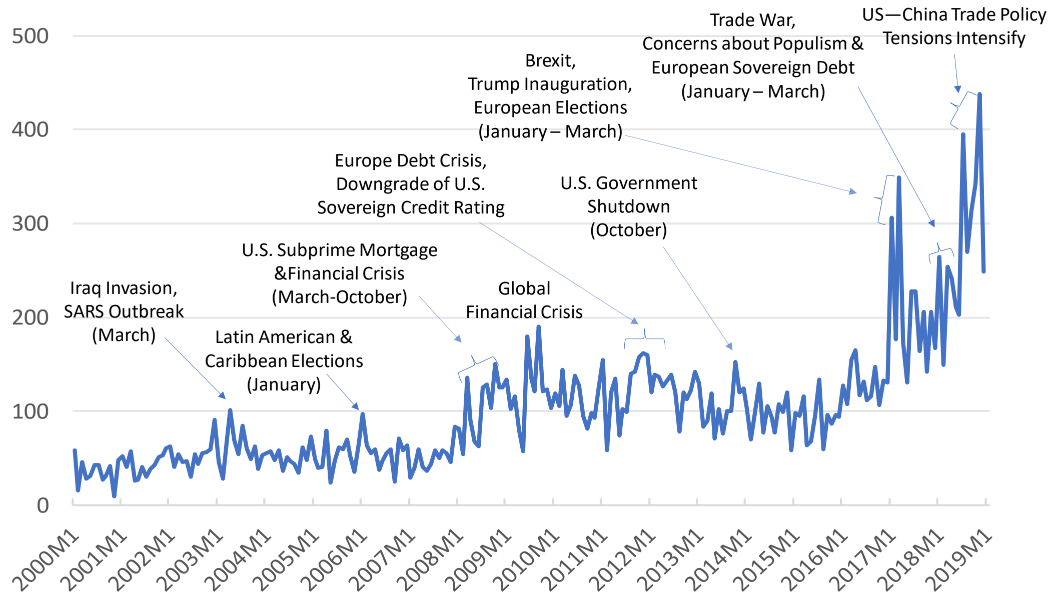

Figure 2 China Monthly EPU Index since 2000

Chinese stock market was established in December 1990 and has been highly impacted by governmental policy. There is a large amount literature pointed out that there's a deviation of the stock market return from economic condition. Therefore, it is worth to survey how the economic policy uncertainty (EPU) affect the stock market, and if it is a relevant predictor of Chinese stock market return.

Different from forecasting other economic variables, stock return has low out-of-sample forecastability, despite extensive in-sample evidence of equity premium predictability. Stock returns inherently contain a sizeable unpredictable component, so that the best forecasting models can explain only a relatively small part of stock return. Traditionally, regression-based methods are used to predict stock return. However, they barely pay attention to model uncertainty and parameter uncertainty. Recently, improved forecasting strategies delivered statistically and economically significant out-of-sample gains, relative to the historical average bench mark. These strategies improve forecasting performance by addressing the substantial model uncertainty and parameter uncertainty surrounding the data-generating process for stock return.

Unlike other economic variables, stock market has higher data generating frequencies. Rational asset pricing theory posits out that stock return predictability can result from exposure to time-varying aggregate risk, and to the extent that successful forecasting models consistently capture this time-varying aggregate risk premium, they will likely remain successful over time. This indicates the importance to a time-variant model with time-variant parameters. DMA models satisfy these two conditions.

Capitals are sensitive to news (new related information). As a result, stock market return behaves higher volatility when good news or bad news come out. The volatilities are a comprehensive result of all the information happen within the observation period. For example, a big negative news would make the daily stock market return has a higher probability to move towards negative.

Moreover, the economic regime or policy tendency changes over time. The famous Lucas critique, "Given that the structure of an economic model consists of optimal decision rules of economic agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any change in policy will systematically alter the structure of economic models". His words point out two difficulties in forecasting work. First, we'll never know what the best model is when we are doing model specification, and the best model might change over the time. Moreover, using too many regressors would potentially bring in over-fitting problem for the model. Second, just as the structure of series changes, the coefficients of each corresponding predictors also change along the time. Hence, a model with time-invariant coefficients usually have poor performance in forecasting, as it takes the data from every point of time equivalently.

On the contrary, Time Varying Parameter (TVP) models allow all the estimated parameters change at each state or each time point. Instead of treating all the data points equivalently, it gives more weights on the data at nearest time points.

This paper uses efficient Dynamic Model Averaging (eDMA). Dynamic Model Averaging (DMA) not only allows for coefficients to change over time, but also allows for the entire forecasting model to change overtime. Furtherly, eDMA allows a grid search of the forgetting factor, shortening estimation time more than MCMC used in DMA. The results in the empirical part can show that the forgetting factor does change over the time. Thus, there would be model misspecification problem if a forecaster ignored model uncertainty, parameter instability and changes in forgetting factor.

The changes in forgetting factor is also intuitively correct. When there is zero or mild underline structure shifts, the forgetting factor should be closer to 1, indicating mild changes between neighbor time points. In comparison, while there is a big policy change in stock market, we expect an inconsistency in the data generating process (DGP). Thus, models with a lower forgetting factor over-perform that with a higher forgetting factor.

2. Econometric Methodology

Traditionally, stock market return is predicted with the following linear regression model with selected variables:

$$r_{t+1} = \alpha + \sum_{k=0}^{p} \varphi_k r_{t-k} + \sum_{k=0}^{q} \theta_k X_{t-k} + \varepsilon_{t+1} \qquad (1)$$

Where r_t is the stock market return and X represents a matrix of other explanatory variables, $\varepsilon_t \sim N(0, H)$.

A simple regression with all relevant variables like equation (1) is usually called “kitchen sink” (KS) model, which has three potential issues:

First, it compromises all potential explanatory variables, causing severe over-fitting problem. Overfitting problem is implausible for out-of-sample forecast.

Second, for samples with a large time range, it is very likely that relevant predictors change over the time. Thus, KS model causes severe model misspecification problem. Bayesian model averaging (BMA) is an application to solve this problem.

$$r_{t+1}^{(m)} = \alpha^{(m)} + \sum_{k=0}^{p} \varphi_k^{(m)} r_{t-k}^{(m)} + \sum_{k=0}^{q} \theta_k^{(m)} X_{t-k}^{(m)} + \varepsilon_{t+1}{}^{(m)} \qquad (2)$$

Where $\varepsilon_t{}^{(m)} \sim N\left(0, H^{(m)}\right)$. In equation (2), there are $K = p + q$ explanatory variables and $M = 2^K$ model candidates. Instead of using a single model, BMA provide a posterior probability to each of the M models. The posterior probability will be used as the weight when averaging the prediction results across all the models. The total number of coefficients to be estimated is $M \times K$, which are impossible to be estimated individually. Koop etc. (2007) suggests using MC^3 to estimate. However, BMA models fail to consider parameter uncertainty. As for each of the M models, BMA assumes the coefficients’ role stays consistent all the time.

Third, the coefficient of each predictors may also be time variant, indicating the marginal effect of each predictors may change over time. Time-varying parameter (TVP) models has the advantages to account for parameter uncertainty.

$$r_{t+1} = \alpha_t + \sum_{k=0}^{p} \varphi_{k,t} r_{t-k} + \sum_{k=0}^{q} \theta_{k,t} X_{t-k} + \varepsilon_{t+1} \qquad (3)$$

$$\begin{pmatrix} \alpha_t \\ \varphi_t \\ \theta_t \end{pmatrix} = \begin{pmatrix} \alpha_{t-1} \\ \varphi_{t-1} \\ \theta_{t-1} \end{pmatrix} + \eta_t \qquad (4)$$

Where φ_t is a vector of all $\varphi_{k,t}$, and θ_t is a vector of all $\theta_{k,t}$, $\varepsilon_t \sim N(0, H_t)$, $\eta_t \sim N(0, Q_t)$, the errors are assumed to be mutually independent at all leads and legs.

However, TVP models fail to consider model uncertainty. Comparing (1) and (3), the coefficients in (3) becomes time variant. If K is large, TVP models still suffer from over-fitting problem, and hence have poor out-of-sample forecasting performance.

DMA model

DMA model considers both model uncertainty and parameter instability. Furthermore, it uses two forgetting factors and Kalman filter to simplify computation. The main model looks as follow:

$$r_{t+1}{}^{(m)} = \alpha_t{}^{(m)} + \sum_{k=0}^{p} \varphi_{k,t}{}^{(m)} r_{t-k}{}^{(m)} + \sum_{k=0}^{q} \theta_{k,t}{}^{(m)} X_{t-k}{}^{(m)} + \varepsilon_{t+1}{}^{(m)} \qquad (5)$$

$$\begin{pmatrix} \alpha_t \\ \varphi_t \\ \theta_t \end{pmatrix}^{(m)} = \begin{pmatrix} \alpha_{t-1} \\ \varphi_{t-1} \\ \theta_{t-1} \end{pmatrix}^{(m)} + \eta_t{}^{(m)} \qquad (6)$$

Where $\varepsilon_t{}^{(m)} \sim N\left(0, H_t{}^{(m)}\right)$, $\eta_t{}^{(m)} \sim N\left(0, Q_t{}^{(m)}\right)$,

Like the TVP model, the errors in DMA models are assumed to be mutually independent at all leads and legs. The DMA model allows for parameter variation at each time point for each model m. By calculating the posterior model probability for individual model m at each time point in time in the form of $P_{r_t}(m|r^{t-1})$. To forecast with DMA model, one can take weighted average with this probability:

$$p(\hat{r}_t|r_{t-1}) = \sum_{m=1}^{M} p(\hat{r}_t|r_{t-1}, m_t)p(m_t|r_{t-1}) \quad (7)$$

Where $p(\hat{r}_t|r_{t-1})$ shows the predictive density, M represents all models used at each time point

In comparison, to forecast with Dynamic Model Selection (DMS), one can select the model with highest probability at each time point.

However, a shortcoming of model (5) and (6) is how to compute the evolution of the model over time, as there are $2^{M\times T}$ parameters to be estimated. DMA solves this problem by introducing two important new forgetting factors: parameter forgetting factor λ, and model forgetting factor α. The key aspect of Raftery et al. (2010) algorithm is to use Kalman filter instead of using MCMC. It assumes each model m at time t, is highly correlated with itself at time $t-1$. Similarly, coefficient ${\varphi_{k,t}}^{(m)}$ is highly correlated with ${\varphi_{k,t-1}}^{(m)}$, and the same with ${\theta_{k,t}}^{(m)}$. The correlation is limited to within model, it doesn't apply across model. For example, ${\varphi_{k,t}}^{(m)}$ is independent with ${\varphi_{k,t}}^{(l)}$ and ${\varphi_{k,t-1}}^{(l)}$, where l stands for any model other than m.

For given values of H_t and Q_t, the standard Kalman filter carries out recursive estimation or forecasting. Typically, Kalman filter begins with:

$$ {\Theta_{t-1}}^{(m)}|r^{t-1} \sim N\left({\widehat{\Theta}_{t-1}}^{(m)}, {\Sigma_{t-1|t-1}}^{(m)}\right) \quad (8)$$

Where $\Theta_{t-1} = \begin{pmatrix} \alpha_{t-1} \\ \varphi_{t-1} \\ \theta_{t-1} \end{pmatrix}$. Equation (8) only depends on H_t and Q_t. Kalman filter processes the recursive estimation by:

$$ {\Theta_t}^{(m)}|r^{t-1} \sim N\left({\widehat{\Theta}_{t-1}}^{(m)}, {\Sigma_{t|t-1}}^{(m)}\right) \quad (9)$$

Where

$$ {\Sigma_{t|t-1}}^{(m)} = {\Sigma_{t-1|t-1}}^{(m)} + {Q_t}^{(m)} \quad (10)$$

To simplify the calculation, Raftery et al. (2010) made an approximation with the use of the parameter forgetting factor λ.

$$ {\Sigma_{t|t-1}}^{(m)} = \lambda^{-1} {\Sigma_{t-1|t-1}}^{(m)} \quad (11)$$

Equivalently, (11) indicates $Q_t{}^{(m)} = (\lambda^{-1} - 1)\Sigma_{t-1|t-1}{}^{(m)}$, where $0 < \lambda \leq 1$. When $\lambda = 1$, $\Sigma_{t|t-1}{}^{(m)} = \Sigma_{t-1|t-1}{}^{(m)}, Q_t{}^{(m)} = 0$. It means that the model does not allow time variation, DMA becomes BMA. In application, it's common to set λ near 1, suggesting a very gradual evolution of coefficients. Raftery et al. (2010) set $\lambda = 0.99$, meaning observations 20 periods earlier only receives 81% as much weight as last period's observations. Comparatively, if $\lambda = 0.95$, observations 20 periods earlier only receive 35% as much weight as last period's observations, implying a rapid evolution of coefficients. Therefore, a small λ indicates a dramatic change on the parameter and thus indicates some structural change or shock happened in the system. Koop et al. (2012) suggested the reasonable values for $\lambda \in (0.95, 0.99)$ and chose $\lambda = 0.99$ for the bench mark choice with an analysis of the sensitivity. In this paper, I'm using $\lambda = [0.90, 0.95\ 0.99]$ for grid search.

This approximation avoids estimation or simulation of $Q_t{}^{(m)}$ for each time t, simplifying the computation for time-variant coefficients. Hence, we can focus on the method to estimate $H_t{}^{(m)}$.

Allowing time variation on the forgetting factor, Leopoldo et al. (2016) proposed effective DMA (eDMA) a grid search of λ_t.

$$\Sigma_{t|t-1}{}^{(m)} = \frac{1}{\lambda_t}\Sigma_{t-1|t-1}{}^{(m)} \qquad (12)$$

When finishing this updating equation (9) becomes:

$$\Theta_t{}^{(m)}|r^t \sim N\left(\widehat{\Theta}_t{}^{(m)}, \Sigma_{t|t}{}^{(m)}\right) \quad (13)$$

where

$$\widehat{\Theta}_{t|t}{}^{(m)} = \widehat{\Theta}_{t|t-1}{}^{(m)} + \Sigma_{t|t-1}{}^{(m)} z_t' \left(H_t{}^{(m)} + z_t \Sigma_{t|t-1}{}^{(m)} z_t'\right)^{-1} \left(r_t - z_t \widehat{\Theta}_{t-1}{}^{(m)}\right) \quad (14)$$

And

$$\Sigma_{t|t}{}^{(m)} = \Sigma_{t|t-1}{}^{(m)} - \Sigma_{t|t-1}{}^{(m)} z_t' \left(H_t{}^{(m)} + z_t \Sigma_{t|t-1}{}^{(m)} z_t'\right)^{-1} z_t \Sigma_{t|t-1} \quad (15)$$

Where $z_t = \begin{pmatrix} 1 \\ r_{t-1} \\ \dots \\ r_{t-k} \\ X_{t-k} \end{pmatrix}$.

Then, recursive forecasting could be done using the predictive distribution:

$$r_t{}^{(m)}|r^{t-1} \sim N\left(z_t \widehat{\Theta}_{t-1}{}^{(m)}, H_t{}^{(m)} + z_t \Sigma_{t|t-1}{}^{(m)} z_t'\right) \quad (16)$$

The all results above were conditional on arbitrary model m. However, we need to consider model transition for each time t, which is for unconditional prediction. For the same reason stated above, if we specify a transition matrix P, such as a Markov transition matrix, it would be computational heavy. Thus, Raftery et al. (2010) involves a model forgetting factor for the state equation for the models, α.

Therefore, the estimation of coefficient follows:

$$p(\Theta_{t-1}|r^{t-1}) = \sum_{m=1}^{M} p\left(\Theta_{t-1}{}^{(m)}\middle|L_{t-1} = m, r^{t-1}\right) \Pr(L_{t-1} = m|r^{t-1}) \quad (17)$$

Where $p\left(\Theta_{t-1}{}^{(m)}\middle|L_{t-1} = m, r^{t-1}\right)$ is given by equation (8). Let $\pi_{t|s,l} = \Pr(L_t = l|r^s)$ to simplify notation. Thus, $\Pr(L_{t-1} = m|r^{t-1}) = \pi_{t-1|t-1,m}$.

The unrestricted model transaction probability from model l to model m is:

$$\pi_{t|t-1,m} = \sum_{l=1}^{M} \pi_{t-1|t-1,l} p_{ml} \quad (18)$$

It requires to calculate for each p_{ml}. To simplify computation, Raftery et al. (2010) replace p_{ml} with model forgetting factor α:

$$\pi_{t|t-1,m} = \frac{\pi^{\alpha}{}_{t-1|t-1,m}}{\sum_{l=1}^{M} \pi^{\alpha}{}_{t-1|t-1,l}} \quad (19)$$

Where $0 < \alpha \leq 1$ is set to a fixed value slightly less than one and is interpreted in a similar manner to λ. This type of approximation was firstly discussed by Smith and Miller (1986). Raftery et al. (2010) agreed with them about it. They argued that it is an empirical sensible simplification and is a type of multiparameter power steady model. This approximation is sensible and not too restrictive.

The most benefit of using this approximation is that it does not require an MCMC or MC^3 algorithm. Similarly, Kalman filter is used to update equation:

$$\pi_{t|t,m} = \frac{\pi_{t|t-1,m} p_m(r_t|r^{t-1})}{\sum_{l=1}^{M} \pi_{t|t-1,l} p_l(r_t|r^{t-1})} \quad (20)$$

Where $p_l(r_t|r^{t-1})$ is the predictive density for model l evaluated at r_t.

Thus, recursive forecasting can be done by averaging over predictive results for every model using $\pi_{t|t-1,m}$. Then, DMA point forecasts are given by:

$$E(r_t|r^{t-1}) = \sum_{m=1}^{M} \pi_{t|t-1,m} z_t^{(m)} \widehat{\Theta}_{t-1}^{(m)} \quad (21)$$

Comparatively, DMS selects the single model with highest probability $\pi_{t|t-1,m}$ at each time t.

The discussing above is all conditional on H_t. Raftery et al. (2010) recommend a simple plug in method where $H_t^{(m)} = H^{(m)}$ and is replace with a consistent estimate. Thus, the fixed H is also the initial input for estimation.

3. Empirical Results

The forecasting is done in real time, with a full sample runs from 1992:M1 to 2018:M5. The in-sample period is 1992:M1-2000:M10, as I chose the first 1/3 sample to initiate the model. The forecast horizons of $h = 1, 3, 6, 12, 24$.

The predictors are:

- SSE: Shanghai Security Exchange Composite Index monthly return, the market index.
- M2: Broad money supply, including cash and checking deposits (M1) as well as near money.
- CPI: Inflation Rate.
- Deposit Rate: the deposit rate regulated by the central bank of China.
- Real GDP: interpolated monthly real GDP from quarterly data.
- Investment: the fixed investment.
- Consumption: residents' consumption.
- EPU: Economic Policy Uncertainty Index.
- Turning Point: economic contraction dummy variable, 1 for the economy is in economic contraction.
- Discount rate: the banking discount rate.
- CCI: consumer confidence index.
- Exr: exchange rate, RMB to USD.
- M1: cash and checking deposits.

I also include 3 lags of SSE as predictors, resulting in 15 predictors in total with an intercept. The model is specified as follows:

$$\begin{aligned} SSE_t = \alpha_t &+ \theta_{1,t} SSE_{t-1} + \theta_{2,t} SSE_{t-2} + \theta_{3,t} SSE_{t-3} + \theta_{4,t} m2_t + \theta_{5,t} cpi_t + \theta_{6,t} deposit\ rate_t \\ &+ \theta_{7,t} realgdp_t + \theta_{8,t} investment_t + \theta_{9,t} consumption_t + \theta_{10,t} epu_t \\ &+ \theta_{11,t} turning\ point_t + \theta_{12,t} discount\ rate_t + \theta_{13,t} cci_t + \theta_{14,t} exr_t + \theta_{15,t} m1_t \\ &+ \varepsilon_t \end{aligned}$$

All input variables are made stationary.

Table 1 Unit Root Test of Variables

	ADF	KPSS	PPerron
sse	-8.21***	0.35	-306.64***
dlm2	-4.77***	2.23***	-377.32***
ddcpi	-10.56***	0.02	-363.78***
dldr	-5.00***	0.1289	-315.86***
dlrealgdp	-6.00***	1.02***	-382.13***
dlinv	-6.56***	0.33	-336.36***
dlcons	-4.58***	0.52**	-339.27***
dlepu	-10.38***	0.02	-381.3***
drec	-3.59**	0.04	-314.82***
dldr1	-5.62***	0.16	-320.76***
dlcci	-6.55***	0.11	-72.12***
dyexr	-3.73**	0.32	-176.69***
dlm1	-5.34***	0.99***	-161.93***

Note: ***, **, * stand for 99%, 95% and 90% significant level respectively.

3.1 Is EPU a useful predictor?

As stated in introduction part, I expect EPU to be a strong predictor for short term forecast $h = 1$. Shown in **Figure 3** (green line) below, the posterior probability of EPU to be included as a predictor stays around 0.6-0.8 persistently. The only predictor that performs higher posterior level is the one lag of the stock market return SSE_{t-1}, shown in **Figure 4** (red line). Surprisingly, monetary policy variables M1, M2, discount rate and deposit rate does not seem to be good predictors for Chines stock market return, implying some theoretical discrepancy, which are shown in figures in Appendix.

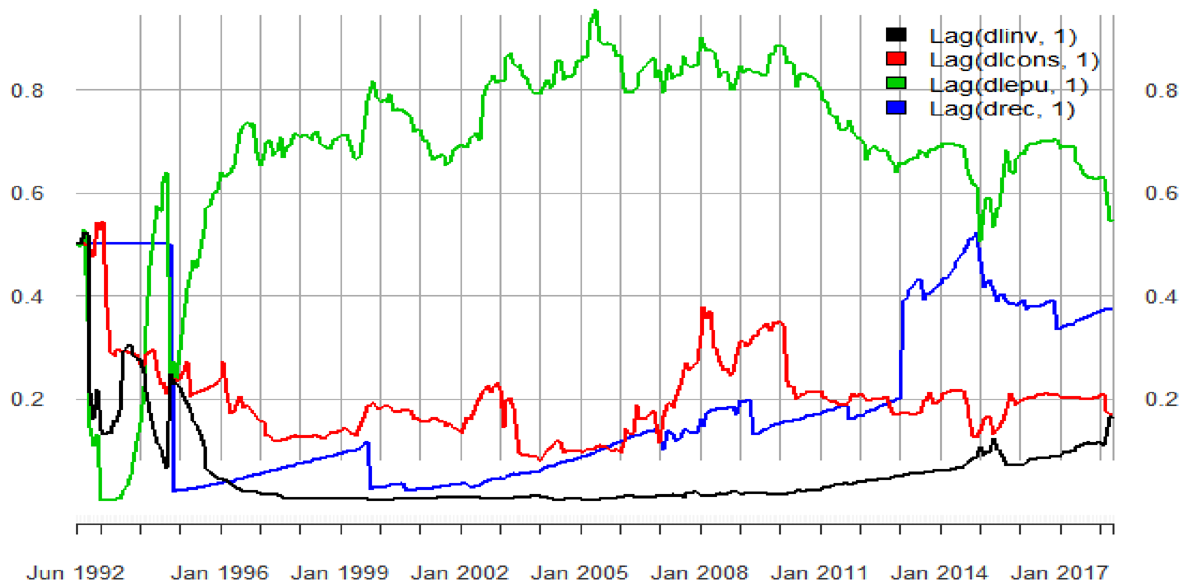

Figure 3 Posterior Inclusion Probability of EPU Compared to Monetary Policy Variables (h=1)

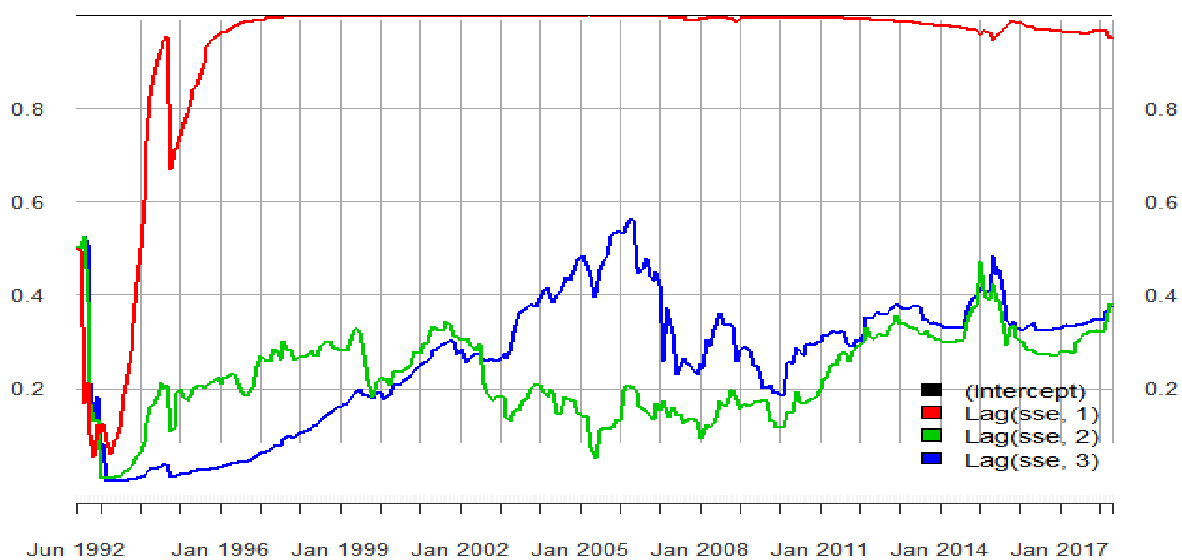

Figure 4 Posterior Inclusion Probability of SSE_{t-1}, SSE_{t-2} *and* SSE_{t-3} *(h=1)*

Furthermore, EPU has stronger prediction power for short-run (h=1) forecast (black line). Figure 5 is a comparison of posterior inclusion probability for EPU of forecast horizon $1, 3, 6, 12, 24$.

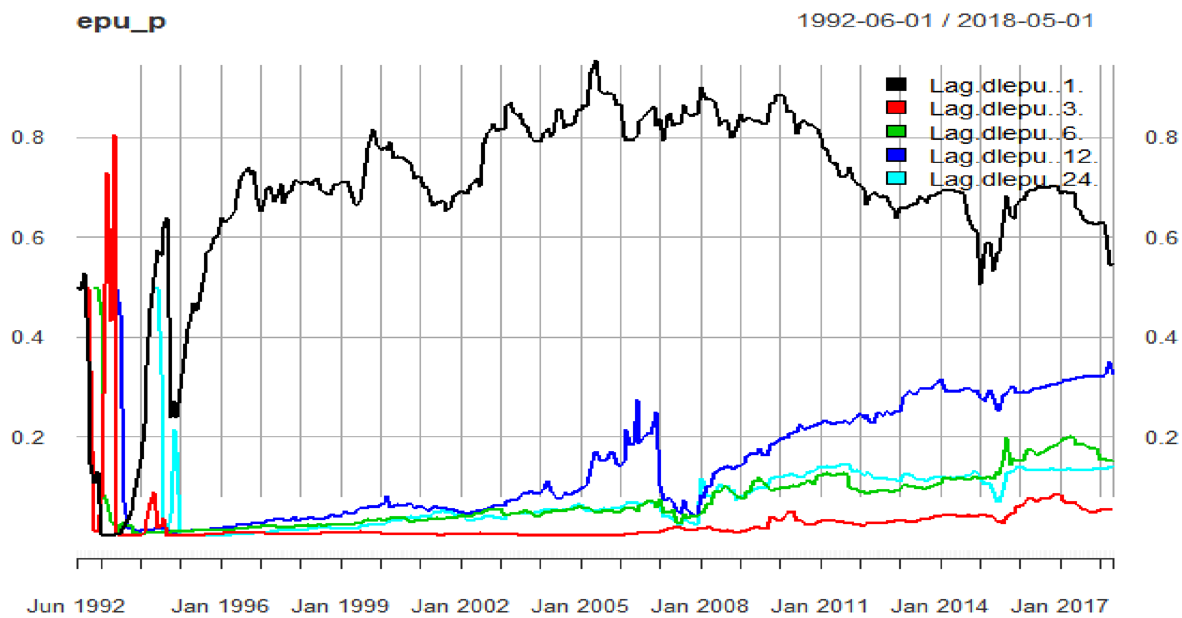

Figure 5 Posterior Inclusion Probability of EPU across Forecast Horizon

3.2 Does eDMA outperform other forecasting methodologies?

This paper uses AR(1) as the bench mark model. I also tested BMA, BMS and eDMS. To further investigate if EPU has explanatory power, eDMA without EPU is considered. Since the out-of-sample forecasting was conducted for the later 2/3 period, the $iBurnPeriod\ = 105$. That is, to evaluate the forecasting for the later 2/3 period. The forecasting evaluation is $\frac{MSE_{model}}{MSE_{baseline}}$, and the baseline model is AR(1). Thus, the smaller the value, the better forecast evaluation the model performs.

Table 2 Forecasting Evaluation

	h=1	h=3	h=6	h=12	h=24
BMA	0.0422	0.3684	1.0000	**1.0000**	**1.0000**
BMS	0.0422	0.3684	1.0000	**1.0000**	**1.0000**
EDMA	**0.0365**	**0.3199**	**0.8571**	**1.0000**	1.0733
EDMS	0.0401	0.3579	1.0000	**1.0000**	1.1283
EDMA_without EPU	0.0366	0.3319	**0.8571**	**1.0000**	1.0733
EDMS_ without EPU	0.0410	0.3684	1.0000	**1.0000**	1.1283

Table 2 shows that eDMA with EPU outperforms all other models. Firstly, it shows that EPU is a strong predictor to stock market return. Secondly, it shows the advantages of eDMA to other forecasting models. There are two main reasons to discuss below.

First, eDMA captures model uncertainty and parameter instability. This can be seen from Figure 1 and other figures in the appendix. Figure 6 also shows that, eDMA picks different predictor combinations for each time t, and hence reduces over-fitting problem. The size stands for the weight-averaged number of predictors in the model.

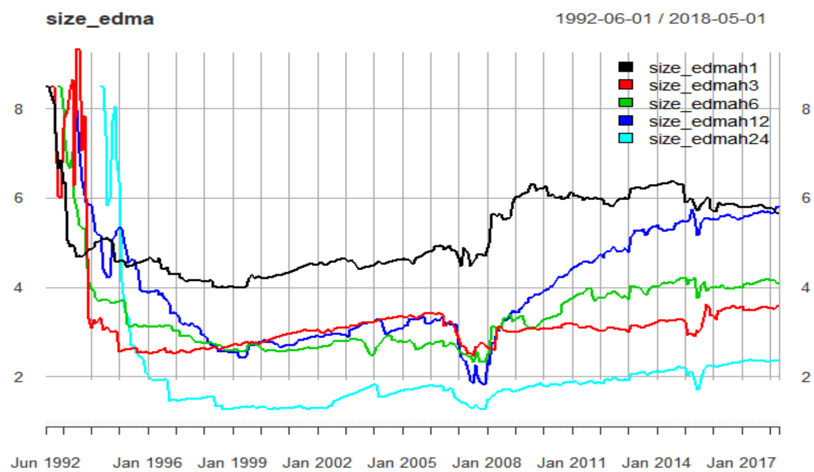

Figure 6 Size of Predictors

Secondly, compared to TVP model, eDMA uses forgetting factor to reduce the massive estimation requirement of state space model. As well known, random walk (RW) outperforms many models and the most important reason is, that RW does not require estimation of parameters. The more estimation we need to make, the more uncertainty we introduce to the model. TVP has this problem and thus performs poorly. eDMA not only simplify the estimation problem, but also allows changes in the parameter forgetting factor λ, which can be seen in Figure 7. It shows that the parameter forgetting factor also changes along the time. It is worth noticing that the forgetting factor drops dramatically during last economic recession, implying structural changes in real economy

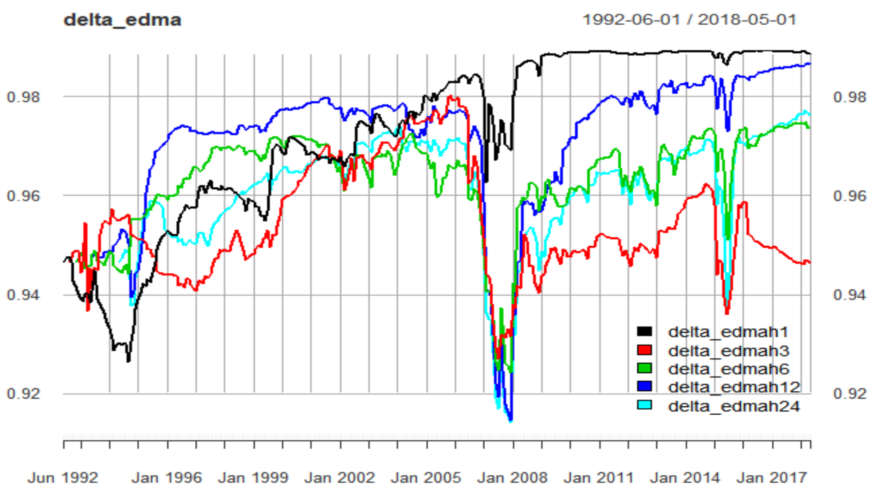

Figure 7 Posterior weighted average of delta

4. Conclusion

This paper shows that EPU has a strong predictive power for Chinese stock market return. Potentially because it captures the underline volatility patter of stock market. Also, EPU acts best for short term (1 month) stock market return forecasting. This means EPU has more potential power for higher frequency stock return forecasting, as the stock market acts fast to new information and the impact of single news does not persist for a long time. Thus, it is necessary to survey higher frequency stock return, for example, weekly and biweekly.

eDMA also outperforms other models investigated in this paper. It outperforms for its consideration of model uncertainty and parameter instability. Besides, it conservatively considers the predictors to include in the model and thus reduce the over-fitting problem. Additionally, it simplifies the computation, which not only is more efficient for application, but also reduces uncertainty in estimation.

Moreover, it’s clear in Figure 4 and 5 that the 2008 financial crisis in the U.S. had a big impact on underline data generation process (DGP) of the Chinese stock market return.

References

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. "Measuring economic policy uncertainty." *The Quarterly Journal of Economics* 131, no. 4 (2016): 1593-1636.

Bloom, Nicholas. "The impact of uncertainty shocks." *Econometrica* 77, no. 3 (2009): 623-685.

Brogaard, Jonathan, and Andrew Detzel. "The asset-pricing implications of government economic policy uncertainty." *Management Science* 61, no. 1 (2015): 3-18.

Catania, Leopoldo, and Nima Nonejad. "Dynamic model averaging for practitioners in economics and finance: The eDMA package." *arXiv preprint arXiv:1606.05656* (2016).

Koop, Gary, Dale J. Poirier, and Justin L. Tobias. *Bayesian econometric methods*. Cambridge University Press, 2007.

Koop, Gary, and Dimitris Korobilis. "Forecasting inflation using dynamic model averaging." *International Economic Review*53, no. 3 (2012): 867-886.

Pastor, Lubos, and Pietro Veronesi. "Uncertainty about government policy and stock prices." *The journal of Finance*67, no. 4 (2012): 1219-1264.

Raftery, Adrian E., Miroslav Kárný, and Pavel Ettler. "Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill." *Technometrics* 52, no. 1 (2010): 52-66.

Steven J. Davis, Dingqian Liu and Xuguang Simon Sheng. "Economic Policy Uncertainty in China Since 1949: The View from Mainland Newspapers", paper in progress

Appendix

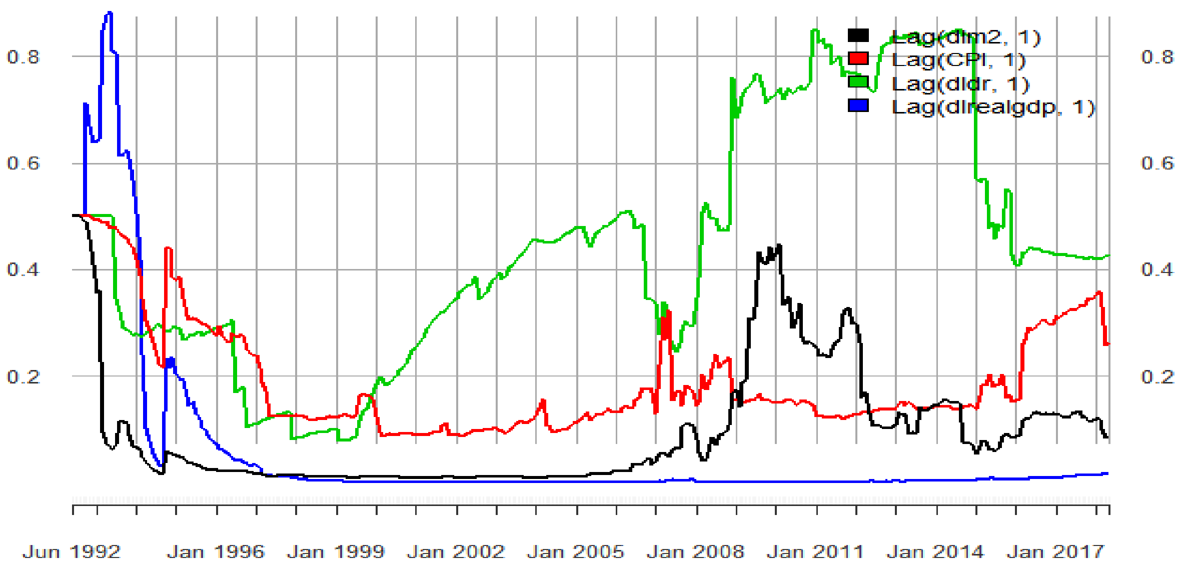

Figure 8 Posterior Inclusion likelihood

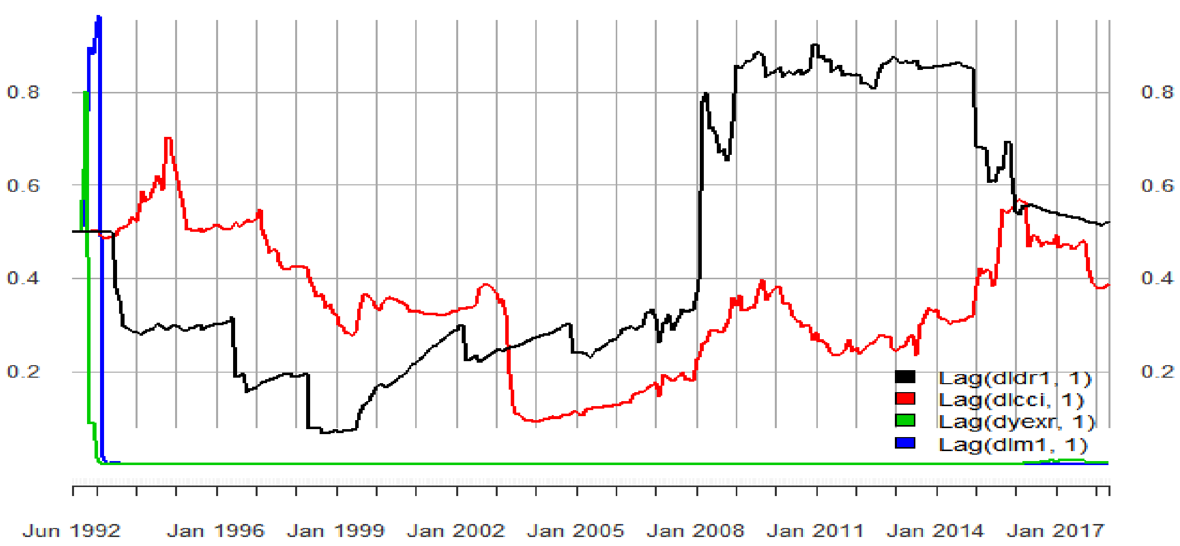

Figure 9 Posterior Inclusion likelihood

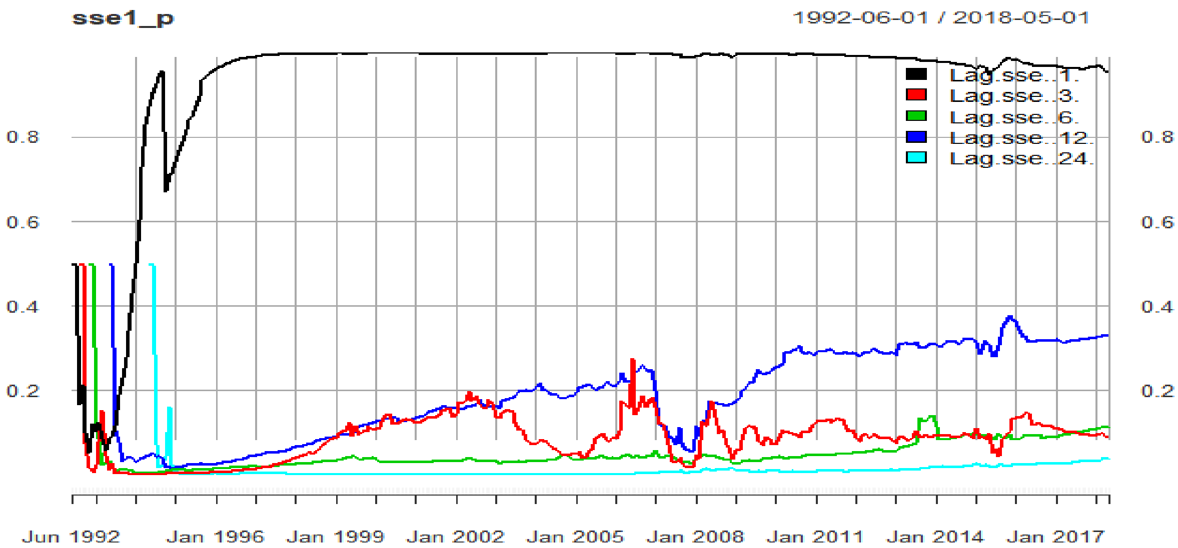

Figure 10 Posterior Inclusion likelihood across forecast horizon

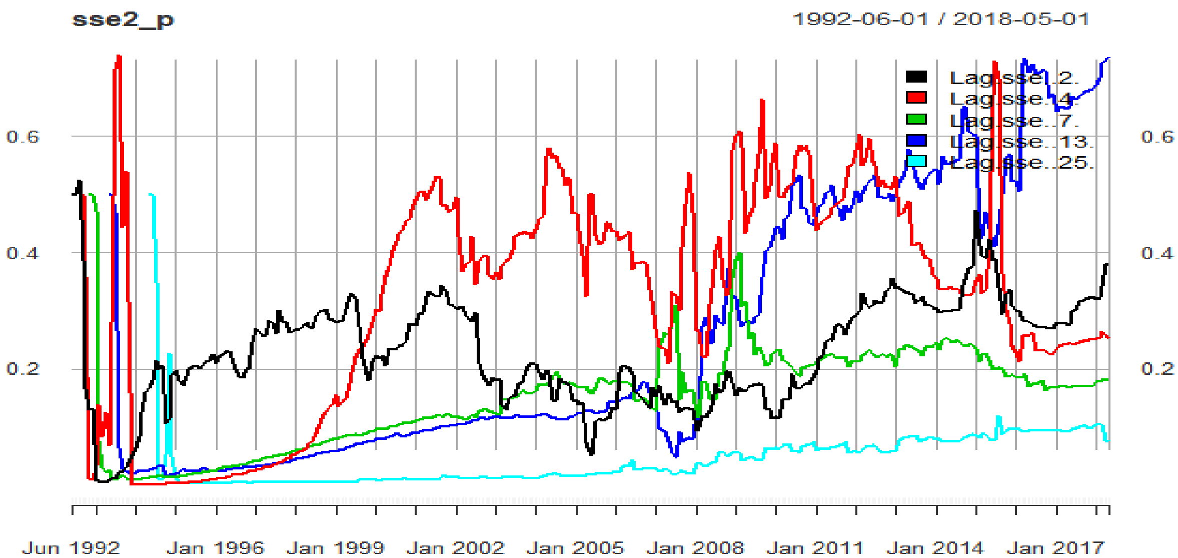

Figure 11 Posterior Inclusion likelihood across forecast horizon

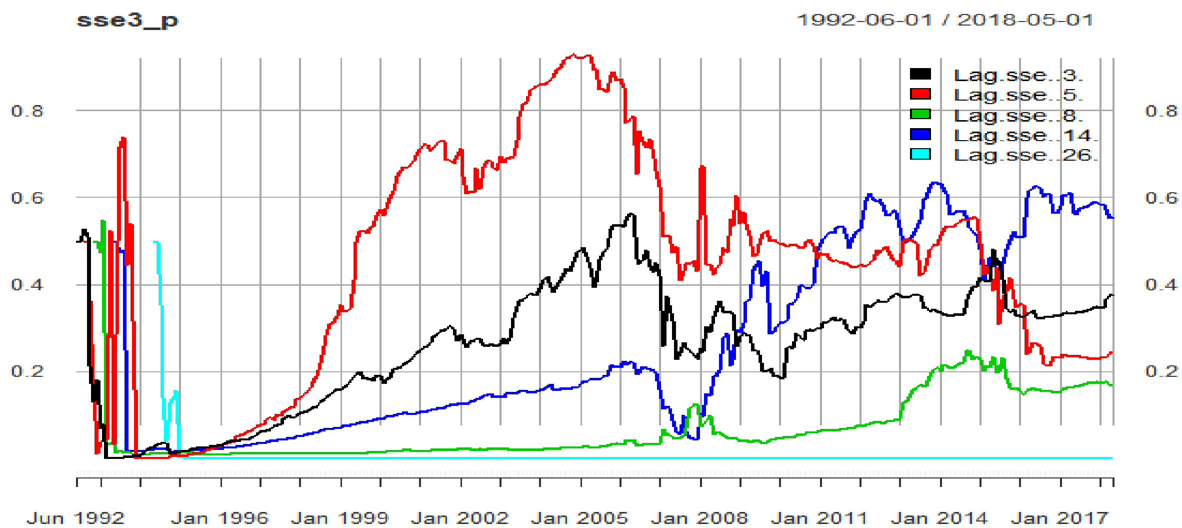

Figure 12 Posterior Inclusion likelihood across forecast horizon

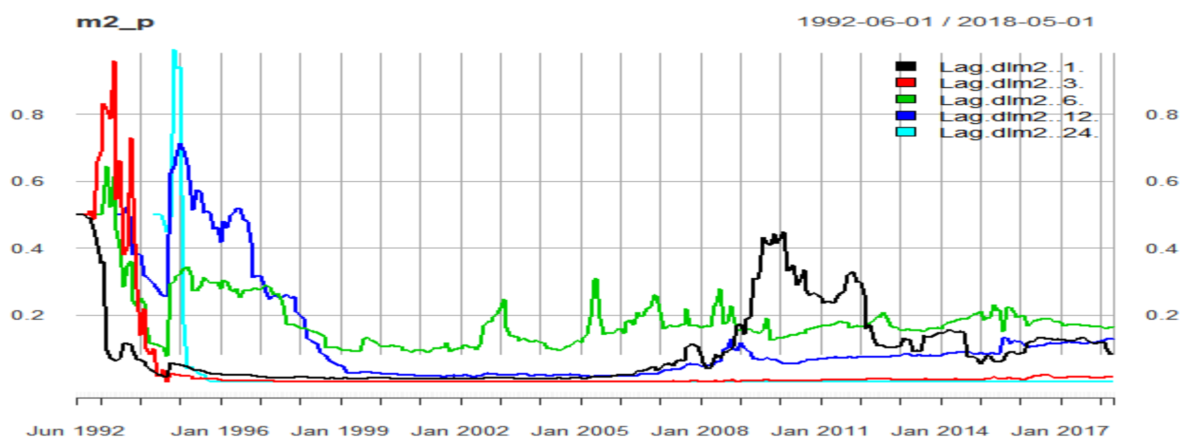

Figure 13 Posterior Inclusion likelihood across forecast horizon

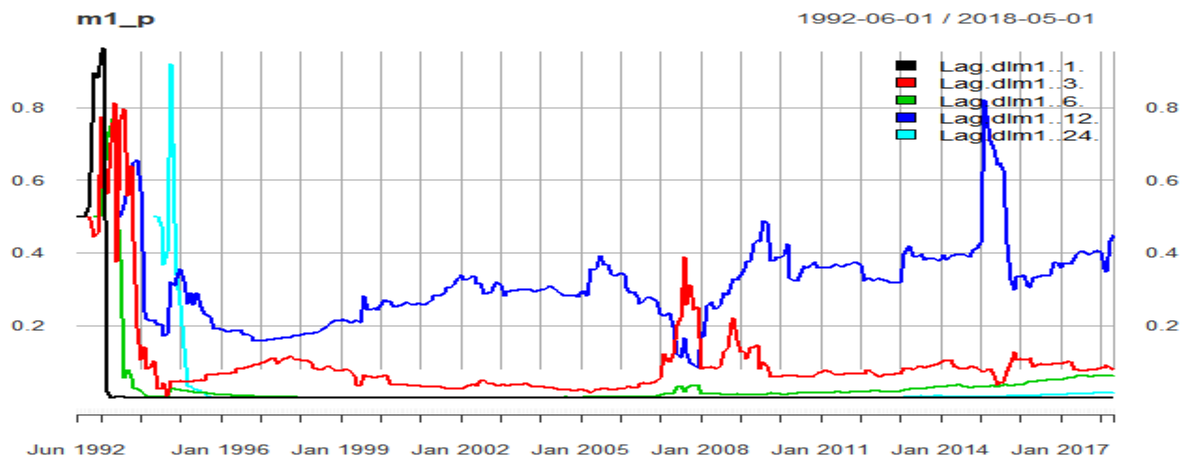

Figure 14 Posterior Inclusion likelihood across forecast horizon

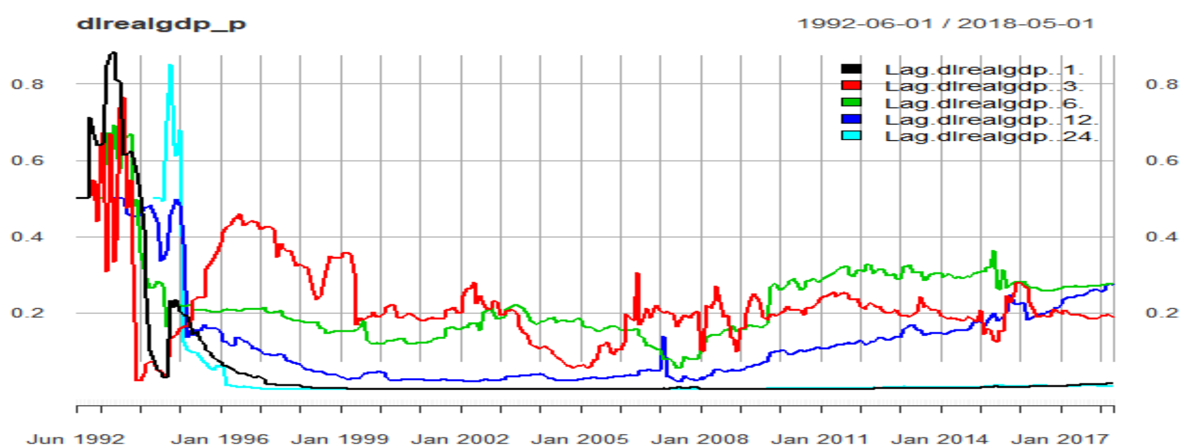

Figure 15 Posterior Inclusion likelihood across forecast horizon

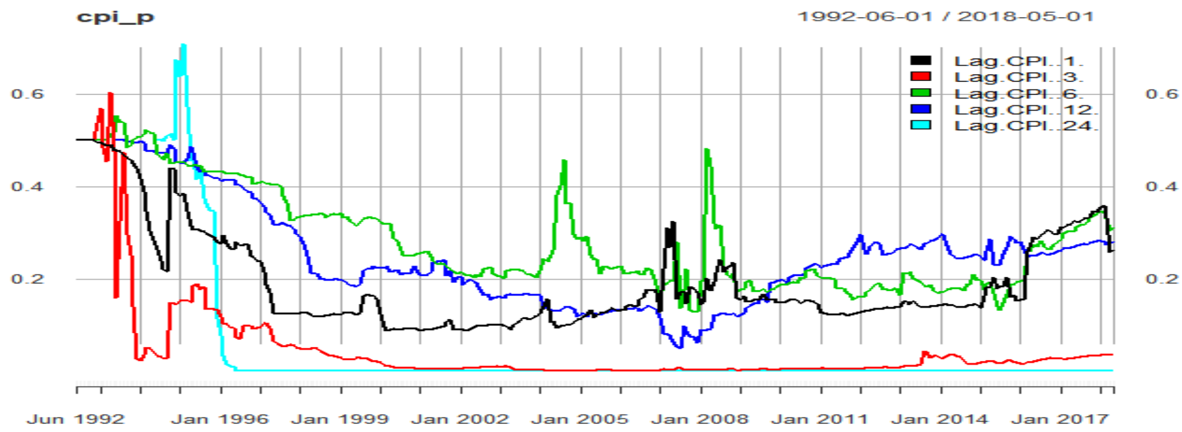

Figure 16 Posterior Inclusion likelihood across forecast horizon

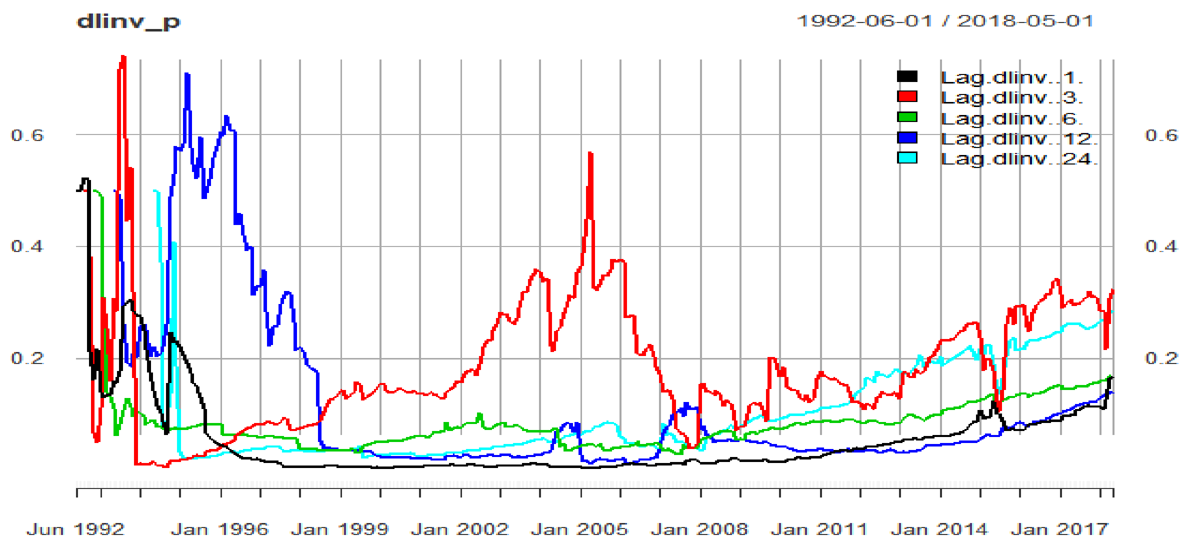

Figure 17 Posterior Inclusion likelihood across forecast horizon

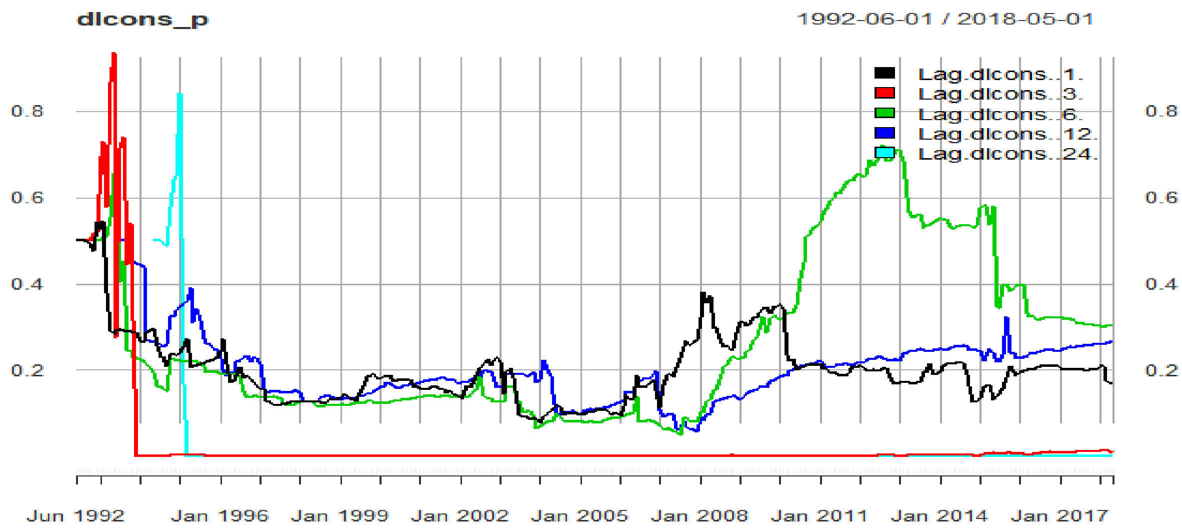

Figure 18 Posterior Inclusion likelihood across forecast horizon

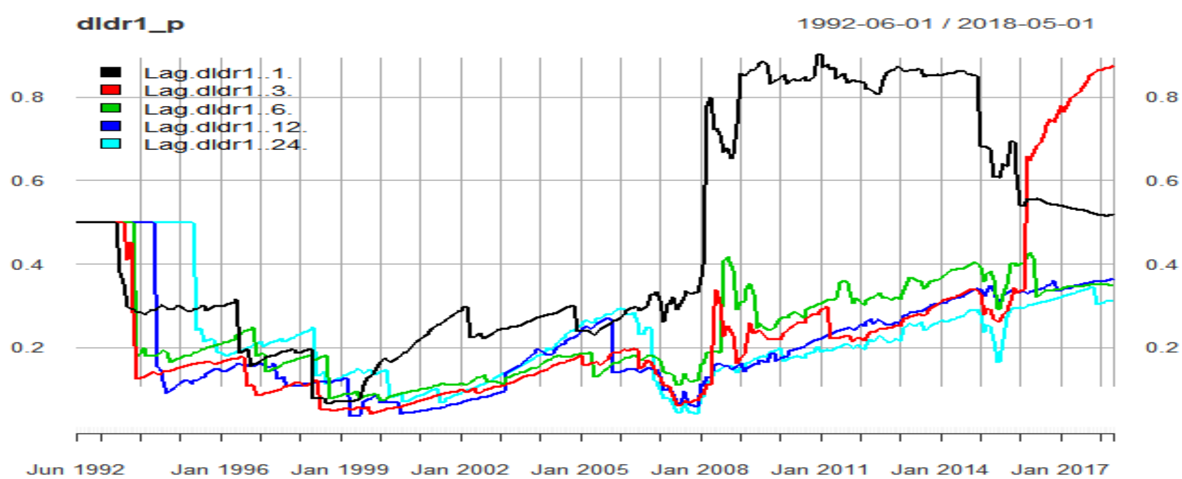

Figure 19 Posterior Inclusion likelihood across forecast horizon

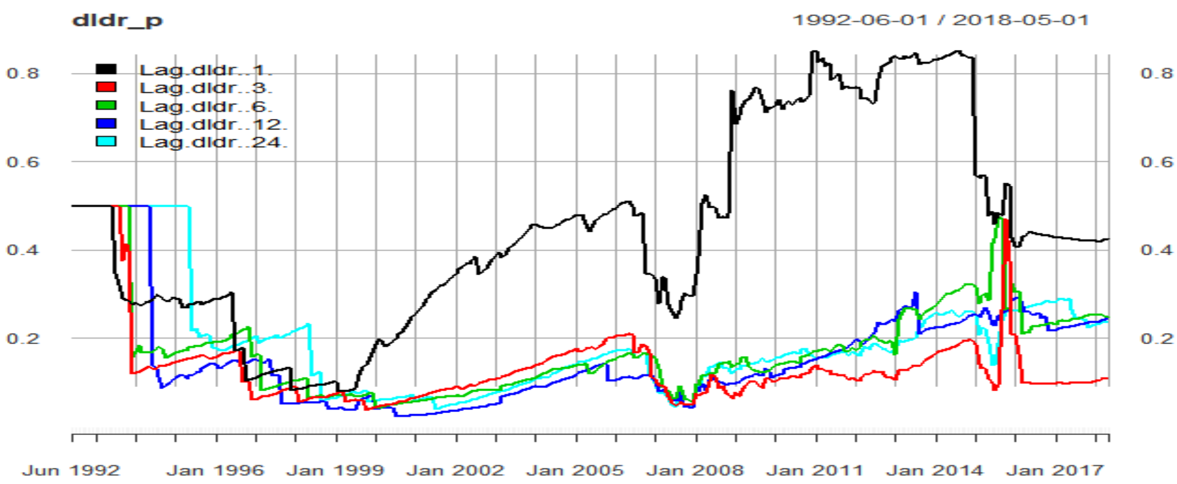

Figure 20 Posterior Inclusion likelihood across forecast horizon

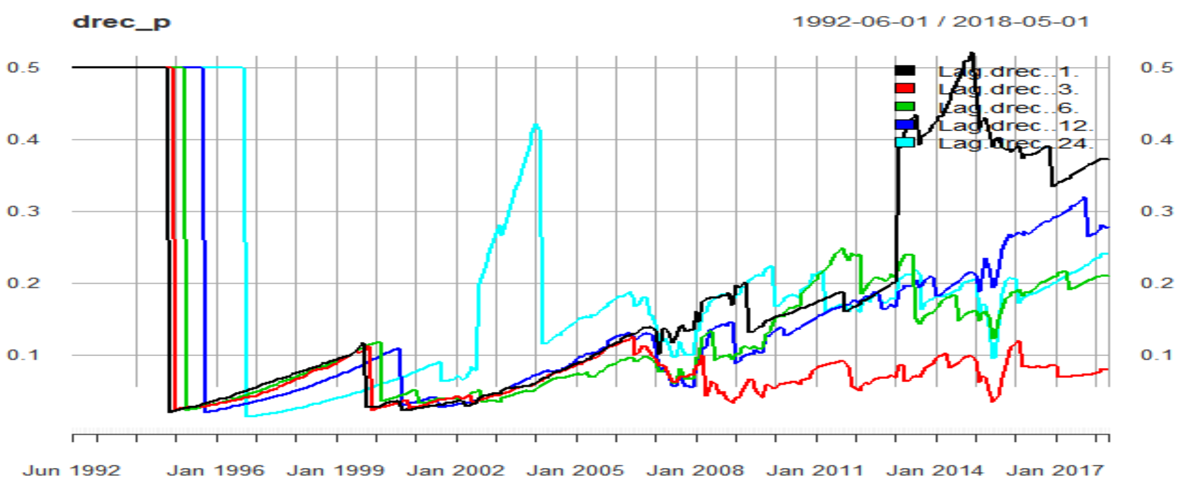

Figure 21 Posterior Inclusion likelihood across forecast horizon

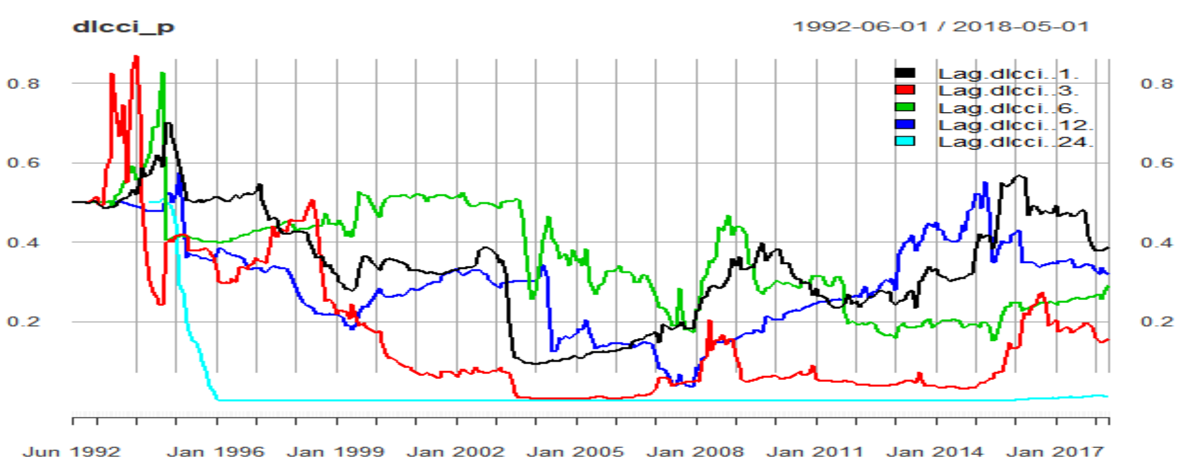

Figure 22 Posterior Inclusion likelihood across forecast horizon

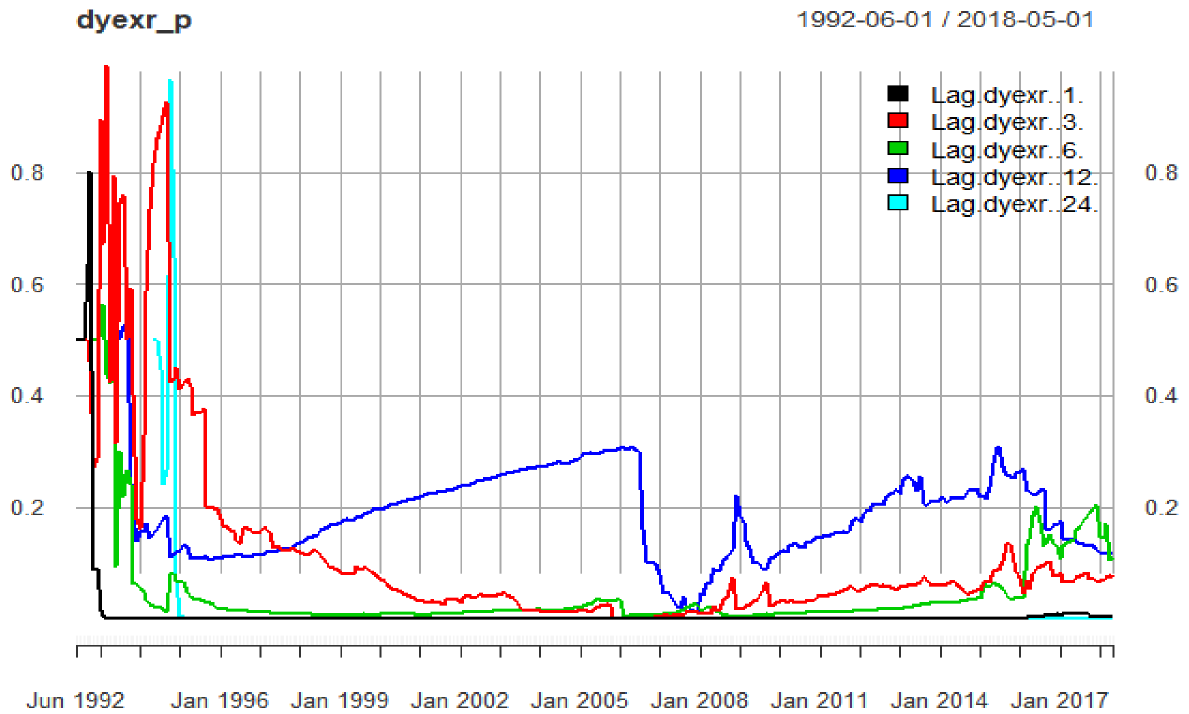

Figure 23 Posterior Inclusion likelihood across forecast horizon