Forecasting the yield curve - Forecast performance of the dynamic Nelson-Siegel model from 1971 to 2008

Molenaars, Tomas K. and Reinerink, Nick H. and Hemminga, Marcus A.

RiskCo BV, Utrecht, The Netherlands, RiskCo BV, Utrecht, The Netherlands, RiskCo BV, Utrecht, The Netherlands and Laboratory of Biophysics, Wageningen University, The Netherlands

11 July 2013

Forecasting the yield curve - Forecast performance of the dynamic Nelson-Siegel model from 1971 to 2008

Tomas K. Molenaars[1], Nick H. Reinerink[1], Marcus A. Hemminga[1,2,*]

[1]RiskCo BV
Financial Product Design and IT
Utrecht
The Netherlands

[2]Laboratory of Biophysics
Wageningen University
Wageningen
The Netherlands

[*]**Corresponding author**
Dr. Marcus A. Hemminga
RiskCo BV
Financial Product Design and IT
Maliebaan 68
3581 CV Utrecht
The Netherlands
Tel: +31 30 2059994
E-mail: marcus.hemminga@riskco.nl
Web site: http://www.riskco.nl/

Running title:
Forecasting the yield curve

JEL classification:
- C5 - Econometric Modeling
- E4 - Macroeconomics and Monetary Economics
- G17 - Financial Forecasting and Simulation

Keywords:
- Term structure of interest rates
- Yield curve modeling
- Dynamic Nelson-Siegel model
- Out-of-sample forecasting evaluations

Version 2, July 11, 2013

ABSTRACT

We define a parameter representing the relative forecast performance to compare forecasting results of different methods. By using this parameter, we analyze the performance of the dynamic Nelson-Siegel model and, for comparison, the first order autoregressive (AR(1)) model applied to a set of US bond yield data that covers a time span from November 1971 to December 2008. As a reference, we take the random walk model applied to the yield data. Our findings indicate that none of the models can convincingly beat the random walk model. Furthermore, there is no advantage in using the more advanced and complicated dynamic Nelson-Siegel model over a simple AR(1) model.

1. Introduction

A yield curve (i.e., the term structure of interest rates) represents the relationship between interest rates and the remaining time to maturity. Forecasting of the yield curve can provide important information for monetary policy, as it provides a basis for investment and saving strategies. In this view, the development of models for forecasting yield curves is of fundamental importance to banks and financial institutions, such as life insurers and pension funds.

Modeling the zero-coupon yield curve has already a long history, starting with the equilibrium models of Vasicek (Vasicek, 1977) and Cox et al. (Cox et al., 1985). These are short rate models by which yield curves can be constructed. Later on it was shown that these models produce poor forecasts, since better forecasts could be generated by assuming that yields follow a random walk model (Duffee, 2002). Alternatively, Ho and Lee (Ho and Lee, 1986), Hull and White (Hull and White, 1990), and Heath et al. (Heath et al., 1992) developed models that focused on fitting the term structure at a given point in time to ensure that no arbitrage possibilities exist, which is important for pricing derivatives. However, such methods are not appropriate for forecasting (Duffee, 2002; Koopman et al., 2010).

An alternative approach was proposed by Diebold and Li (Diebold and Li, 2006). They constructed forecasting models based on the Nelson-Siegel model (Nelson and Siegel, 1987) and tested the forecast performance using US Treasuries bond yields. This dynamic Nelson-Siegel model (Christensen et al., 2009) utilizes a set of exponential components whose contributions are analyzed as a function of time. This method, in fact, is based on modeling the yield curve using its shape. Diebold and Li argued that their approach forecasts well, especially for a 6 and 12-month forecast horizon. Coroneo et al. (Coroneo et al., 2011) showed that the Nelson-Siegel model in a statistical sense is compatible with the no-arbitrage constraints on the US market. The forecasting performance of a number of extensions of the Nelson-Siegel model was examined by De Pooter (De Pooter, 2007), suggesting that a model that adds a second slope to the Nelson-Siegel model forecasts particularly well. Recently, Chen and Niu (Chen and Niu 2013) proposed an extension of the dynamic Nelson-Siegel model that adaptively detects parameter changes and forecast the yield curve, showing a good 3- to 12-months ahead forecast performance.

In the studies with the dynamic Nelson-Siegel model (Coroneo et al., 2011; De Pooter, 2007; Diebold and Li, 2006) limited attention was paid to systematically examine the forecast performance of the model. Therefore, in this paper we examine the dynamic Nelson-Siegel model using the US Treasuries bond yields for a historic data set ranging from November 1971 to December 2008. We evaluate a large number of forecasting results of the dynamic Nelson-Siegel and, for comparison, the first order autoregressive (AR(1)) model. As a reference, we take a random walk model applied to the yield data.

2. Data

Similar to Diebold and Li (Diebold and Li, 2006), we use end-of-month price quotes (bid-ask average) for US Treasuries from CRSP government bonds files. Those bond prices are converted to unsmoothed Fama-Bliss forward rates (Fama and Bliss, 1987), which are further transformed to unsmoothed Fama-Bliss zero-coupon yields. This data set is provided by Robert Bliss and covers the period from November 1971 (1971:11) to December 2008 (2008:12) with maturities 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months.

3. Theory and methodology

3.1. Modeling approach

In this paper, we analyze the yields by using three different models: a random walk model, an AR(1) model and the dynamic Nelson-Siegel model.

The random walk model is given by:

$$y_t(\tau) = y_{t-1}(\tau) + \epsilon_t(\tau), \qquad (1)$$

where *y* is the zero-coupon yield, *t* the time (i.e., the date of the yield curve) in months, τ the time to maturity in months and $\epsilon_t(\tau)$ is a white noise process.

For AR(1) on the yields, we have:

$$y_t(\tau) = a(\tau) + b(\tau) y_{t-1}(\tau) + \epsilon_t(\tau), \qquad (2)$$

where $a(\tau)$ and $b(\tau)$ are constants per maturity τ. These parameters are obtained by regression on the time series of the yields.

In the case of the dynamic Nelson-Siegel model, the yield curve is fitted with the following Nelson-Siegel equation (Diebold and Li, 2006):

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{3,t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right). \qquad (3)$$

In this equation we have four time-dependent parameters, which can be interpreted as follows: the shape parameter λ_t governs the exponential decay rate and parameters $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ represent the contribution of the so-called long-term component, short-term component and medium-term component, respectively. Eq. (3) is not linear in λ_t, hence for every time *t* we should estimate the parameters by a nonlinear fit. However, we follow the approach of Diebold and Li (Diebold and Li, 2006), by fixing $\lambda_t = \lambda$. This avoids potentially challenging numerical optimizations. Doing this enables us to estimate the remaining parameters $\beta_{i,t}$ by ordinary least-squares regression. The resulting times series for these parameters are modeled subsequently using the AR(1) model according to Eq. (2).

3.2.Forecasting procedures

The models that we use in the forecasting procedures described above are summarized in Table 1. The forecast by the random walk model (RW) is given by:

$$\hat{y}_{t+h}(\tau) = y_t(\tau), \qquad (4)$$

where $\hat{y}_{t+h}(\tau)$ is the forecast of the yield, and *h* is the forecast horizon. Thus this forecast in fact is a no-change forecast.

For the AR(1) model on the yields, the forecast is:

$$\hat{y}_{t+h}(\tau) = \hat{a}_h(\tau) + \hat{b}_h(\tau) y_t(\tau). \quad (5)$$

Here $\hat{a}_h(\tau)$ and $\hat{b}_h(\tau)$ are the estimated parameters. We abbreviate this model by AR.

In the forecasting procedures with the dynamic Nelson-Siegel model (NS) in Eq. (3), the AR(1) forecast for the parameters $\beta_{i,t}, i = 1,2,3,$ can be written as:

$$\hat{\beta}_{i,t+h} = \hat{a}_{i,h} + \hat{b}_{i,h}\beta_{i,t}, \quad (6)$$

where $\hat{a}_{i,h}$ and $\hat{b}_{i,h}$ are the estimated parameters. Assuming a constant value for λ, the forecasted yield curve at time *t*+*h* is given by:

$$\hat{y}_{t+h}(\tau) = \hat{\beta}_{1,t+h} + \hat{\beta}_{2,t+h}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \hat{\beta}_{3,t+h}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right). \quad (7)$$

To evaluate the out-of-sample performance of a forecasting procedure, we calculate the root-mean-square-error (RMSE), given by

$$\mathrm{RMSE}_{model}(\tau) = \sqrt{\frac{1}{T-t_0}\sum_{t=t_0}^{T}\left(\hat{y}_t(\tau) - y_t(\tau)\right)^2}, \quad (8)$$

where $\hat{y}_t(\tau)$ is the forecasted yield of the model, $y_t(\tau)$ is the yield from the data, and $[t_0, T]$ is the interval of times for which we make the forecasts. The smaller the RMSE, the better the forecast quality of the model.

To be able to compare the quality of the forecast results of the models, we define the relative forecast performance *F*, which is the relative difference in forecast error of the model with respect to the RW model:

$$F_{model} = \frac{\sum_\tau \mathrm{RMSE}_{RW}(\tau) - \sum_\tau \mathrm{RMSE}_{model}(\tau)}{\sum_\tau \mathrm{RMSE}_{RW}(\tau)}, \quad (9)$$

where $\sum_\tau \mathrm{RMSE}_{RW}(\tau)$ and $\sum_\tau \mathrm{RMSE}_{model}(\tau)$ sum over the RMSE values at all maturities τ of the random walk model and fitting model, respectively.

We take the random walk model as our bench mark, as it has the most simple no-change forecast, to provide a minimum standard on predictive accuracy for each model. Positive values of *F* denote a better forecast of the model as compared to the random walk model; negative values indicate a reduced performance. By definition, the relative forecast performance of the random walk model is 0.

Our forecasting approach is as follows. We estimate the models in Table 1 by using an expanding window starting with 108 months. We make predictions for 84 months for all maturities as specified in Section 2. We forecast ahead for five different forecast horizons *h*: 1, 3, 6, 9, and 12 months with a fixed shape parameter $\lambda = 0.0609$ (in month^{-1}), following Diebold and Li (Diebold and Li, 2006) [1]. The period sensitivity is examined by changing the

[1] In this paper, Diebold and Li argue that the value of λ_t that maximizes the medium-term component in Eq. (3) at exactly 30 months is $\lambda_t = 0.0609$. This statement is not correct. The medium-term component has a bump shape with a maximum at $\lambda_t\tau = 1.793$. From this relationship, it can be seen that $\lambda_t = 0.0609$ actually corresponds to 29.44 months.

starting date, and therefore by shifting the data set for analysis and forecasting. In addition, we vary parameter λ systematically to examine its effect on the 6-months ahead forecasts.

3.3. Modeling and forecasting algorithms

Curve fitting is carried out by using ordinary least squares. The curve fitting, data analysis methods and forecasting procedures were constructed in MATLAB (MathWorks, Inc., Natick, MA, USA).

4. Forecasting results

Our forecasting results are presented in Fig. 1A and B, which show the relative forecast performance F of the models NS and AR, respectively, as a function of time at forecast horizons of 1, 3, 6, 9, and 12 months. In this figure on the horizontal axis, the starting dates are shown for the various forecast periods. For example, 1994 (see arrow) reflects the forecast study carried out by Diebold and Li (Diebold and Li, 2006). This point indicates a forecast period from January 1994 up to and including December 2000 (from 1994:1 to 2000:12, i.e., 84 months). For each month in this period the prediction is compared with the real yield curve. To predict the yield curve, the time series parameters are estimated using the time series from a certain window in time, given by the following procedure: to forecast the yield for month m in the period 1994:1 to 2000:12, the months 1985:1 – h to $m - h$ are used for parameter estimation, where h is the forecast horizon. For other starting dates, we use precisely the same yield data set and forecast methods, but the whole set of windows is shifted by months forward and backwards. It should be realized that although we present a single date for each point in Fig.1, it actually reflects a whole series of monthly data that are involved in the underlying calculations.

If we take a look at the relative forecast performance of the AR model F_{AR} in Fig. 1B, it is striking that only for the forecasts between 1993 and 1995 positive values are found, indicating a better performance of the AR model as compared to the RW model. Also positive values for F_{AR} are found from 1983 to 1984, however, this is less convincing. Especially around 1987 the relative forecast performance is strongly negative.

The relative forecast performance of the NS model F_{NS} shown in Fig. 1A follows the trend of F_{AR}. At all dates the relative forecast performance is negative, except for the period from 1993 to 1995. In most cases F_{NS} is worse than F_{AR}. Also it turns out that in most cases for negative values of F_{NS} and F_{AR} the performance reduces with increasing forecast horizon; for positive values of F_{NS} and F_{AR} this effect reverses.

In Fig. 1A, we have taken $\lambda = 0.0609$, which is the value used by Diebold and Li (Diebold and Li, 2006). This value is related to the selection of yield data used by these authors. Since the forecasting result of the dynamic Nelson-Siegel model depends on λ, its effect on F_{NS} is investigated for different values of λ for a forecast horizon of 6 months. This result is presented in Fig. 2, again showing the starting dates for the various forecast periods along the horizontal axis. As can be seen, taking other values for λ does not make much difference, except for $\lambda = 0.03$, which delivers poor forecasts in most cases.

5. Discussion

The main purpose of this paper is to examine the forecasting results of the dynamic Nelson-Siegel model (Eq. (3)) and the first order autoregressive (AR(1)) model, taking the random walk model as a bench mark. We apply these models to the US Treasuries yield data, i.e., the unsmoothed Fama-Bliss zero-coupon yield data (Fama and Bliss, 1987), covering the period from 1971:11 up to 2008:12. In this analysis, we follow exactly the same approach as discussed Diebold and Li (Diebold and Li, 2006).

The models that we use are summarized in Table 1. The dynamic Nelson-Siegel model was introduced by Diebold and Li (Diebold and Li, 2006) to predict yield curves. The AR(1) model is one of the simplest and most widely used autoregressive models. Finally, the random walk model supposes that the prediction of tomorrow's value is today's value. We decided not to examine the vector autoregressive VAR counterpart of the AR(1) model, because the results of Diebold and Li indicate that it performs relatively poor.

To analyze the forecasting performance of the models in Table 1, we calculate the relative forecast performance F (Eq. (9)). This parameter is the relative difference in forecast error of the model with respect to the random walk model, which is taken as a reference. The advantage of using the relative forecast performance F is that it enables us to compare the forecasting results of different models applied to a large yield data set. However, the effect of different maturity values τ is lost in this approach.

Although the maturity details are lost, the relative forecast performance offers an excellent way to analyze the overall trends in the forecasts at different forecast horizons, as can be seen in Fig. 1. In this figure, the monthly relative forecast performance F for the models AR and NS is plotted for a fixed value of λ of 0.0609. The arrow at 1994 represents the relative forecast performance for different forecast horizons for the months 1994:1 to 2000:12 as published by Diebold and Li (Diebold and Li, 2006). Their paper is based on a sample of monthly zero-coupon yield data from 1985:1 to 2000:12. All other results in Fig. 1 are calculated on the same basis, using the full available zero-coupon yield data set, thereby following precisely the approach by Diebold and Li. It should be noted that all underling data used to calculate the relative forecast performance at different values of the forecast horizon at 1994 are exactly in agreement with the results published by Diebold and Li.

In comparing F_{NS} and F_{AR} in Fig. 1, it can be seen that only for 20% of the monthly data points between 1982 and 2002 F_{NS} performs better than F_{AR}. This suggests that there is no convincing advantage in using the more advanced and complicated dynamic Nelson-Siegel model over a simple AR(1) model. This can be understood, because there are a couple of weaknesses in using the dynamic Nelson-Siegel model.

Firstly, one can argue that the Nelson-Siegel curve (Eq. (3)) does not properly fit the yield curve at all dates (for a fixed value of λ). In fact, in the Nelson-Siegel model imposes a functional form to the yield curve. If the yield curve does not fit to this form, the Nelson-Siegel model will result in inferior forecasts. It is well known that adding a fourth term to the Nelson-Siegel equation (the Svensson extension (Svensson, 1995)), which allows for a second "hump/trough", delivers a better yield curve fit. Although there is no fundamental economic theory that supports this Nelson-Siegel-Svensson equation, it is extensively used by Central Banks (BIS, 2005; Gilli et al., 2010). Conversely, in the four-term Nelson-Siegel-Svensson equation more parameters need to be fitted, increasing the risk of fitting noise arising from parameter correlation and multiple local optima (Gilli et al., 2010; Hawkins, 2004).

Secondly, in the estimation of the β-parameters, it is assumed that λ is fixed. This restriction is made to avoid numerical problems with the fit procedures, because we can use ordinary least-squares regression to estimate the β-parameters. However, it is questionable whether the

Nelson-Siegel equation with a fixed λ will properly perform well in all cases. In Fig. 1, we have used a constant value of λ of 0.0609 that is optimized by Diebold and Li (Diebold and Li, 2006) for the result at 1994. The findings in Fig. 2 reveal that the effect of varying λ is small, thus the value of λ will not affect the main conclusions obtained from Fig. 1. Nevertheless, the assumption of a fixed λ may be a source for the low overall relative forecast performance of the dynamic Nelson-Siegel model as compared to the forecast performance of the AR(1) model.

For both parameters F_{NS} and F_{AR} in Fig. 1, a strong dip is observed with a minimum around 1987, indicating a bad performance (up to 40% for a 12-month forecast horizon) for both models NS and AR as compared to the RW model. Interestingly, this date coincides with the so-called Black Monday financial crisis. This indicates that even the AR(1) model, which is a simple extension of the random walk model, already fails to forecast the yield curve in a crisis period better than the random walk model.

The most striking point in Fig. 1, however, is that for most monthly data points the relative forecast performance F is negative, indicating that none of the models AR and NS can convincingly beat the random walk model. Only for a limited time period between 1993 and 1995, the models AR and NS show positive values for the relative forecast performance F, indicating a better performance of the model as compared to the RW model. This leads to the conclusion that, under the selected sample windows and forecast conditions used, neither model is able to convincingly outperform the simple random walk model. This means that the conclusion by Diebold and Li (Diebold and Li, 2006) that the dynamic Nelson-Siegel model, under the forecast setup used, produces 12-months-ahead forecasts that are much more accurate than standard benchmarks is only valid for a very limited time period. This result is supported by Moench (Moench, 2008), who argues that the strong forecast performance of the Nelson-Siegel model documented by Diebold and Li (Diebold and Li, 2006) is partly due to their choice of forecast period. By looking over a long time interval the economic movements are actually random. Thus imposing a model means adding constraints and knowledge that are not actually present in the data. This also explains the fact that at some time periods the Nelson-Siegel model is in favor, but at other time periods it performs bad.

6. Conclusion

In this paper, we analyze the relative forecast performance of the dynamic Nelson-Siegel model and the AR(1) model applied to the zero-coupon US yield data. As a reference, we take the random walk model. Our findings indicate that none of the models can convincingly beat the random walk model, contrary to what is argued by Diebold and Li (Diebold and Li, 2006). Furthermore, there is no advantage in using the more advanced and complicated dynamic Nelson-Siegel model over a simple AR(1) model.

Acknowledgments

We thank Dr. Robert Bliss for kindly providing the US yield data and Dr. Michiel de Pooter for helpful discussions.

References

BIS, 2005, Zero-coupon yield curves: technical documentation. Bank for International Settlements, Basel, Switzerland, BIS Papers 25.

Chen, Y. and Niu, L. (2013). Adaptive dynamic Nelson-Siegel term structure model with applications. SSRN, *http://dx.doi.org/10.2139/ssrn.2025853*.

Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D., 2009, An arbitrage-free generalized Nelson-Siegel term structure model. Econometrics Journal 12, C33-C64.

Coroneo, L., Nyholm, K., Vidova-Koleva, R., 2011, How arbitrage-free is the Nelson-Siegel model? Journal of Empirical Finance 18, 393-407.

Cox, J.C., Ingersoll, J.E., Jr., Ross, S.A., 1985, A theory of the term structure of interest rates. Econometrica 53, 385-407.

De Pooter, M., 2007, Examining the Nelson-Siegel class of term structure models - In-sample fit versus out-of-sample forecasting performance. Tinbergen Institute, Discussion Paper 2007-043/4.

Diebold, F.X., Li, C., 2006, Forecasting the term structure of government bond yields. Journal of Econometrics 130, 337-364.

Duffee, G.R., 2002, Term premia and interest rate forecasts in affine models. The Journal of Finance 57, 405-443.

Fama, E.F., Bliss, R.R., 1987, The information in long-maturity forward rates. American Economic Review 77, 680-692.

Gilli, M., Grosse, S., Schumann, E., 2010, Calibrating the Nelson-Siegel-Svensson model. COMISEF, Working Paper WPS-031.

Hawkins, D.M., 2004, The problem of overfitting. J. Chem. Inf. Comput. Sci. 44, 1-12.

Heath, D., Jarrow, R., Morton, A., 1992, Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica 60, 77-105.

Ho, T.S.Y., Lee, S.-B., 1986, Term structure movements and pricing interest rate contingent claims. The Journal of Finance 41, 1011-1029.

Hull, J., White, A., 1990, Pricing interest rate derivative securities. Review of Financial Studies 3, 573-592.

Koopman, S.J., Mallee, M.I.P., Van der Wel, M., 2010, Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters. Journal of Business and Economic Statistics 28, 329-343.

Moench, E., 2008, Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. Journal of Econometrics 146, 26-43.

Nelson, C.R., Siegel, A.F., 1987, Parsimonious modeling of yield curves. The Journal of Business 60, 473-489.

Svensson, L.E.O., 1995, Estimating forward interest rates with the extended Nelson and Siegel method. Quarterly Review, Sveriges Riksbank 3, 13-26.

Vasicek, O., 1977, An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177-188.

Table 1

Models used in the forecasting procedures. The random walk model (RW) and first order autoregressive (AR(1)) model (AR) are applied directly to the yield data. In the dynamic Nelson-Siegel (NS) model (Eq. (3)), the AR(1) model is applied to the β-parameters from the yield curve fit. In comparing the different forecasting procedures, the random walk model is taken as a bench mark.

Abbreviation	Model type
RW	Random walk model on the yield data
AR	AR(1) model on the yield data
NS	Dynamic Nelson-Siegel model, Eq. (3) and AR(1) on the β-parameters

Figure legends

Fig. 1. Relative forecast performance F of the models NS (A) and AR (B) (see Table 1) for forecast horizons 1, 3, 6, 9, and 12 months (see Section 4). Parameter λ is fixed at a value of 0.0609. The arrow reflects the results of the forecast study carried out by Diebold and Li (Diebold and Li, 2006).

Fig. 2. Effect of λ on the relative forecast performance F of the NS model for a forecast horizon of 6 months (see Section 4).

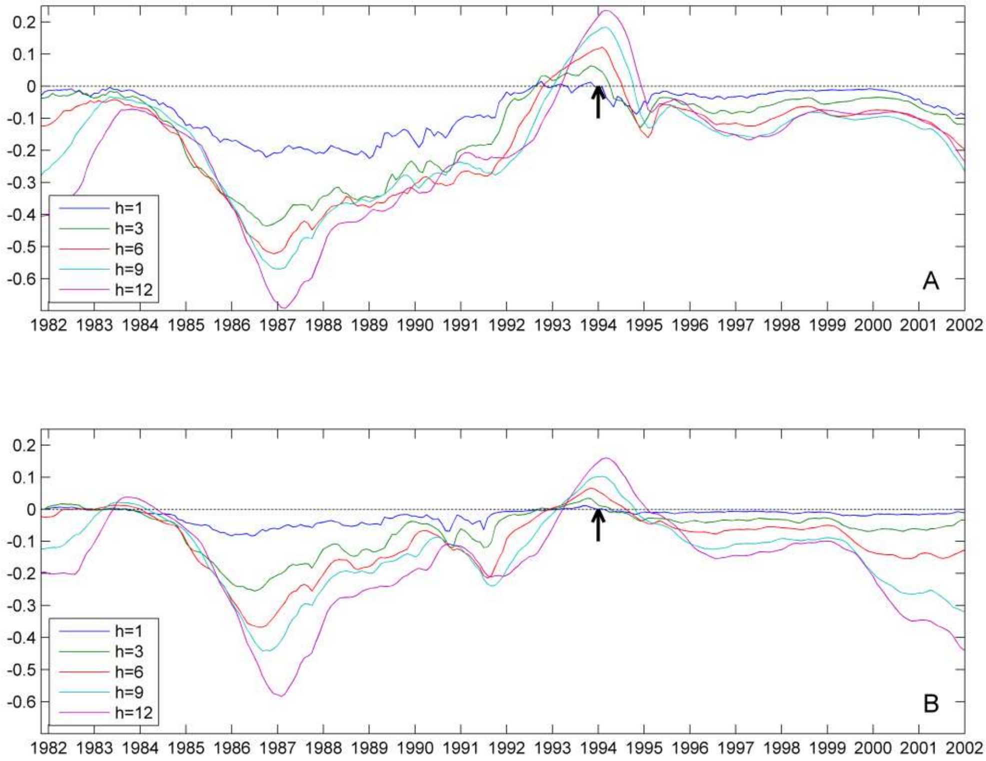

Fig. 1

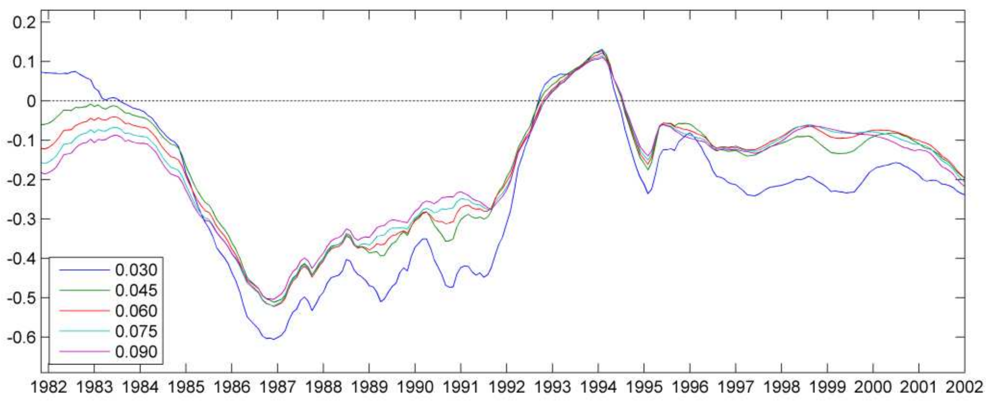

Fig. 2