sustainability

MDPI

Article

Multi-Step Inflation Prediction with Functional Coefficient Autoregressive Model

Man Wang [1], Kun Chen [2,*], Qin Luo [3] and Chao Cheng [4]

1 Department of Finance, Donghua University, Shanghai 200051, China; wangman@dhu.edu.cn
2 School of Statistics, Southwestern University of Finance and Economics, Chengdu 611130, China
3 Guangxi Xijiang Venure Investment Co. Ltd., Nanning 530022, China; Luoqin__1123@163.com
4 Department of Mathmatical Sciences, Tsinghua University, Beijing 100084, China; cqplus@126.com
* Correspondence: chenkun@swufe.edu.cn; Tel.: +86-028-8709-2207

Received: 11 April 2018; Accepted: 16 May 2018; Published: 23 May 2018

Abstract: Forecasting inflation rate is one of the most important topics in finance and economics. In recent years, China has stepped into a "New Normal" stage of economic development, with a different state from the fast growth period during the past few decades. Hence, forecasting the inflation rate of China with a time-varying model may give high accuracy. In this paper, we investigate the problem of forecasting the inflation rate with a functional coefficient autoregressive (FAR) model, which allows the coefficient to change over time. We compare the FAR model based on the B-splines estimation method with the autoregressive moving average (ARMA) model by extensive simulation studies. In addition, with the monthly CPI data of China, we conduct both in-sample analysis and out-of-sample forecasting. The forecasting result shows that the FAR model based on the B-splines estimation method has a better performance than the ARMA model.

Keywords: B-splines; inflation forecast; monthly CPI data; out-of-sample forecast

1. Introduction

The inflation rate is a key index which is closely related to the economic stability and general well-being of a country. It guides policy-makers to formulate the country's macroeconomic and monetary policies. In addition, the households and businesses can make well-informed decisions based on future prices, and investors can construct long-run portfolios based on inflation rate (Bampinas and Panagiotidis, 2016 [1]). Due to these reasons, inflation forecasting attracts much interests from various fields.

The main methods to predict inflation rate include the Phillips curve model, the vector autoregressive-type (VAR) model, and the univariate linear autoregressive moving average (ARMA) model. In the past few decades, although the Phillips curve model has been widely adopted, many research results show that this model cannot provide satisfactory inflation forecasting for countries like China (Stock and Waston, 1999 [2]; Atkeson and Ohanian, 2001 [3]; Mcnelis and Mcadam, 2004 [4]; Matheson, 2006 [5]). On the other hand, the VAR-type model depends on some exogenous factors, for example the real GDP, unemployment, industrial production, manufacturing production, and capacity utilization. Based on this model, inflation forecasting is highly affected by the selection of exogenous factors (see Sekine (2001) [6], Ramakrishman and Vamvakidis (2002) [7] and Ang et al. (2007) [8]). Besides, if the dynamics of the inflation is non-linear, its prediction can be conducted in the framework of non-linear multivariate models. For instance, the regime-switching smooth transition vector autoregressive model used in Lekkos et al. (2007) [9] and the non-linear (asymmetric and polynomial) error correction models used in Milas et al. (2004) [10]. With the development of time series, univariate linear ARMA models without exogenous factors are widely adopted (Bos et al., 2001 [11]; Ang et al., 2007 [8]). In particular, Stock and Waston (1999) [2] found that the univariate linear ARMA models perform better than the Phillips curve model and the VAR-type model in forecasting inflation rate.

However, the univariate linear ARMA model implies that the dynamic mechanism of the underlying process is time-invariant, which is not satisfied by many real time series data (Tong, 1990 [12]). In fact, many empirical studies showed that the dynamics of the inflation rate is time-varying (Chen et al., 2016 [13]).

If a linear ARMA model is used to fit a non-linear data generating process (DGP), the order of the ARMA model is always quite large, resulting in difficult and inaccurate estimation. Therefore, many non-linear time series models are proposed to fit this type of data, for example the threshold autoregressive model of Tong (1990) [12], the bilinear model of Granger and Andersen (1978) [14] and the exponential autoregressive model of Haggan and Ozaki (1981) [15]. These well-known non-linear models belong to parametric models, meaning that one needs to specify the formulation in advance, which is difficult and questionable in many real applications. To overcome this hurdle, some non-parametric time series models are proposed. For example, the non-parametric autoregressive conditional heteroscedastic model and the non-parametric autoregressive model (see Fan and Yao, 2003 [16]).

Although the non-parametric approach is appealing, its application usually requires an unrealistically large sample size when more variables are introduced into the model. This problem is called "curse of dimensionality". In order to avoid this problem and preserve the appreciable flexibility, semi-parametric models are proposed by imposing parametric structures to part of the non-parametric model. Chen and Tsay (1993) [17] proposed the functional-coefficient autoregressive (FAR) model for time series. Similarly, Hastie and Tibshirani (1993) [18] proposed the varying-coefficient model to increase the flexibility of ordinary linear regression model and improve the out-of-sample prediction. Chen and Hong (2012) [19] constructed a test of the smooth transition autoregressive model versus the FAR model. More recently, Chen et al. (2016) [13] established a functional coefficient moving average (FMA) model, with the coefficients estimated by using the local linear estimation technique. Application is made to the monthly CPI data of China. In this paper, we forecast the inflation rate in China by the FAR model, which belongs to the semi-parametric non-linear time series models. The FAR model has three main advantages. First, its formulation is analog to the linear ARMA model and preserve the satisfactory forecasting power. Second, the FAR model is flexible in coefficient specification and easy to interpret. Third, the semi-parametric formulation avoids the possible model mis-specification problem.

A key step to use the FAR model is estimating the time-varying coefficient. The main estimation methods include the kernel estimation method, the local polynomial estimation method, the spline estimation method and the wavelet method. The first two estimation methods belong to local estimation methods, while the last two methods belong to the global estimation method. Based on an iteration algorithm, Chen and Tsay (1993) [17] adopted the arranged local regression estimation and achieved good fitting results. Cai et al. (2000) [20] proposed the local linear estimators and investigated the bandwidth selection problem. Chen and Liu (2001) [21] mainly focused on the local polynomial estimators and associated hypothesis test. Huang and Shen (2004) [22] extend the FAR model to the general functional coefficient regression model and adopted the polynomial spline estimator to estimate the coefficient functions. They showed the consistency of the spline estimator and forecasted the Dutch guilder-US dollar exchange rate based on the estimated model. In this paper, we adopt the B-spline estimators (belonging to the polynomial spline estimator) to estimate the coefficient functions. Compared to other polynomial splines, the B-spline method is numerically stable and the calculation can be obtained recursively, thus significantly reducing the computational burden.

The rest of the paper is organized as follows. Section 2 introduces the FAR model and the spline-based nonparametric estimation method, and provides some implementation details. Section 3 conducts some simulation studies to compare the performance of the FAR model and the ARMA model for fitting data generated from different DGPs. In Section 4, we apply the estimated FAR model to forecast the inflation rate of China. Finally, in Section 5, we give some conclusions.

2. Functional Coefficient Autoregressive Model

2.1. Model Specification

The FAR(p) model proposed by Chen and Tsay (1993) [17] extends the autoregressive (AR) model with the form

$$x_t = \theta_1(U_t)x_{t-1} + \theta_2(U_t)x_{t-2} + \cdots + \theta_p(U_t)x_{t-p} + \varepsilon_t,\ t = 1, \ldots, T, \tag{1}$$

where p is a positive integer representing the lag order, $\{\varepsilon_t\}_{t=1}^T$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance σ^2, $\mathrm{Cov}(x_s, \varepsilon_t) = 0$ for $s < t$, and $\theta_i(\cdot)$ $i = 1, 2, \ldots, p$ are unknown measurable functions. Here, x_t is called the significant variable and U_t is called the threshold variable which can be an exogenous or endogenous variable. In the time series setting, U_t is usually set to be the lagged value of x_t, i.e., $U_t = x_{t-d}$ with $1 \leq d \leq p$. Then the resulting model (denoted as FAR(p, d)) becomes

$$x_t = \theta_1(x_{t-d})x_{t-1} + \theta_2(x_{t-d})x_{t-2} + \cdots + \theta_p(x_{t-d})x_{t-p} + \varepsilon_t,\ t = 1, \ldots, T. \tag{2}$$

Throughout this paper, we impose some conditions on model (1).

Regularity Conditions (RC)

(a) The density function of U_t is nonzero and bounded.
(b) Let $X_t = (x_{t-1}, x_{t-2}, \ldots, x_{t-p})'$, then for any $u \in \mathbb{R}$, the eigenvalues of $\mathrm{E}(X_t X_t' | U_t = u)$ are nonzero and bounded.
(c) The density function of ε_t is positive everywhere.

If **RC** is satisfied, the FAR(p) model (1) is geometric ergodic (Chen and Tsay, 1993 [17]).

2.2. Model Estimation and Prediction Based on B-Spline

We first discuss the identification of the coefficient functions $\theta_i(\cdot), i = 1, \ldots, p$ for model (1). The coefficient functions are said to be identifiable if $\sum_{j=1}^p \theta_j^{(1)}(U_t)x_{t-j} \equiv \sum_{j=1}^p \theta_j^{(2)}(U_t)x_{t-j}$ implies that $\theta_j^{(1)}(U_t) = \theta_j^{(2)}(U_t), j = 1, 2, \ldots, p$. We can prove that the coefficients in model (1) are identifiable under Condition (b). In fact, if we denote $\boldsymbol{\theta}(U_t) = (\theta_1(U_t), \ldots, \theta_p(U_t))'$, then

$$\mathrm{E}\left[\left\{\sum_{j=1}^p \theta_j(U_t)x_{t-j}\right\}^2 \mid U_t = u\right] = \boldsymbol{\theta}(U_t)'\mathrm{E}(X_t X_t' | U_t = u)\boldsymbol{\theta}(U_t). \tag{3}$$

If $\sum_{j=1}^p \left(\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t)\right) x_{t-j} \equiv 0$, then $\mathrm{E}\left[\left\{\sum_{j=1}^p \left(\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t)\right) x_{t-j}\right\}^2\right] = 0$. From this, we can immediately obtain $\mathrm{E}\left[\left\{\sum_{j=1}^p \left(\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t)\right) x_{t-j}\right\}^2 \mid U_t = u\right] = 0$ for any u. By applying (3) and Condition (b), we further get that $\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t) = 0$ almost surely for $j = 1, \ldots, p$.

Since the coefficient functions in model (1) are identifiable, we can now estimate the model. In the literature, the main estimation methods include the kernel estimation method, the local polynomial method, the wavelet method and the spline method, see Fan and Yao (2003) [16] for more details. In particular, as a global smoothing method, the spline method outperforms the kernel estimation method and the local polynomial method for multi-step-ahead forecasting. Moreover, the spline-based method is computationally feasible and the estimated model has a parsimonious explicit expression.

Therefore, we can easily produce multi-step-ahead forecasts by iteratively generating one-step-ahead forecast based on the previous forecasts. On the contrary, if the functional coefficients are estimated based on the local polynomial method, then it is computationally intensive to conduct multi-step-ahead forecasting. Pointed out by Huang and Shen (2004) [22], one needs extra effort to relieve the computational burden for forecasting based on the local polynomial method.

With the aforementioned advantages, the spline-based estimation method is adopted in this paper to estimate the FAR model. Among many different types of splines functions, we adopt the polynomial splines, which are piecewise smooth. This means that the polynomial pieces join together smoothly at a set of interior knot points. Specifically, a polynomial spline of degree $K \geq 0$ on interval $[a, b]$ with knots $a = \xi_0 < \xi_1 < \cdots < \xi_{M+1} = b$ is a polynomial function of degree K on each of the intervals $[\xi_i, \xi_{i+1}), 0 \leq i \leq M-1$ and $[\xi_M, \xi_{M+1}]$, and has $K-1$ continuous derivatives on $[a, b]$.

When $K = 0, 1, 2, 3$, the polynomial splines are called constant spline, linear spline, quadratic spline and cubic spline, respectively. For a given sequence of knots and degree K, the corresponding collection of spline functions form a linear function space. Discussed in de Boor (1972) [23], there are different basis for this space, among which the B-spline basis is most widely used. The B-splines are numerically more stable than other polynomial splines and can be obtained recursively (see Fan and Yao, 2003 [16]). Therefore, we use B-spline basis in this paper for its good numerical property.

In the following, we display the calculation of B-splines. Any spline function of degree K with interior knot sequence $\{\xi_i\}_{i=1}^{M}$ can be expressed as a linear combination of the corresponding B-spline basis $B_{i,K}(x)$, given by

$$B(x) = \sum_{i=0}^{M+K} \beta_i B_{i,K}(x)\,, \tag{4}$$

where $B(x)$ is the spline function and $B_{i,K}(x), i = 0, \ldots, M+K$ are the B-splines. To calculate $B_{i,K}(x)$, we first define $\xi_{-K} = \cdots = \xi_0 < \xi_1 < \cdots < \xi_{M+1}$ as the new augmented sequence of knots, and relabel it as $\eta_0 = \cdots = \eta_K < \eta_{K+1} < \cdots < \eta_{M+K+1}$. The B-splines basis $B_{i,j}(x)$ for $j = 0, 1, \ldots, K$, $i = 0, 1, 2, \ldots, M+K$ are as follows:

$$B_{i,0}(x) = \begin{cases} 1, & \eta_i \leq x \leq \eta_{i+1}, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

and

$$B_{i,j+1}(x) = \frac{x - \eta_i}{\eta_{i+j+1} - \eta_i} B_{i,j}(x) + \frac{\eta_{i+j+2} - x}{\eta_{i+j+2} - \eta_{i+1}} B_{i+1,j}(x). \tag{6}$$

The recursive Formula (6) shows that the B-spline basis relies on the knot sequence and the degree K.

Now, we give an example to demonstrate the calculation process for the B-spline function. Assume the interval is $[0, 1]$ with the interior knot sequence $\{0.25, 0.5, 0.75\}$ and $K = 3$, then the augmented knots vector is $(0, 0, 0, 0, 0.25, 0.5, 0.75, 1)$. When $x = 0.1$, the values of $B_{i,j}(x)$, $i = 0, 1, 2, 3, j = 0, 1, 2, 3$ are shown in Table 1.

Table 1. An example of the calculation process of B-spline.

$j = 3$	$B_{0,3}(0.1)$	$B_{1,3}(0.1)$	$B_{2,3}(0.1)$	$B_{3,3}(0.1)$	$B_{4,3}(0.1)$	$B_{5,3}(0.1)$	$B_{6,3}(0.1)$
	0.216	0.592	0.1813	0.0107	0	0	0
$j = 2$	$B_{0,2}(0.1)$	$B_{1,2}(0.1)$	$B_{2,2}(0.1)$	$B_{3,2}(0.1)$	$B_{4,2}(0.1)$	$B_{5,2}(0.1)$	$B_{6,2}(0.1)$
	0	0.36	0.56	0.08	0	0	0
$j = 1$	$B_{0,1}(0.1)$	$B_{1,1}(0.1)$	$B_{2,1}(0.1)$	$B_{3,1}(0.1)$	$B_{4,1}(0.1)$	$B_{5,1}(0.1)$	$B_{6,1}(0.1)$
	0	0	0.6	0.4	0	0	0
$j = 0$	$B_{0,0}(0.1)$	$B_{1,0}(0.1)$	$B_{2,0}(0.1)$	$B_{3,0}(0.1)$	$B_{4,0}(0.1)$	$B_{5,0}(0.1)$	$B_{6,0}(0.1)$
	0	0	0	1	0	0	0

The B-spline is then used to approximate the coefficient functions $\theta_i(U_t), i = 1, \ldots, p$. According to de Boor (1978) [23] and Schumaker (2007) [24], if $\theta_i(U_t)$ is assumed to be smooth, then as the number of knots tends to infinity, $\theta_i(U_t)$ will be well approximated by a linear combination of the corresponding K-degree B-spline basis. That is, there exist a vector of constants $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{ip})'$ and spline function $\theta_i^\star(U_t)$, such that

$$\theta_i(U_t) \approx \theta_i^\star(U_t) = \sum_{j=0}^{M+K} \beta_{ij} B_{j,K}(U_t),\ i = 1, \ldots, p. \tag{7}$$

Then we approximate model (1) by

$$x_t \approx \sum_{i=1}^{p} \left(\sum_{j=0}^{M+K} \beta_{ij} B_{j,K}(U_t) \right) x_{t-i} + \varepsilon_t\,. \tag{8}$$

Based on (8), the estimation of approximation model (8) is equivalent to the estimation of the vector of parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_p')'$. We estimate $\boldsymbol{\beta}$ by the method of ordinary least squares (OLS), i.e.,

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \sum_{t=p+1}^{T} \left(x_t - \sum_{i=1}^{p} \sum_{j=0}^{M+K} \beta_{ij} B_{j,K}(U_t) x_{t-i} \right)^2 . \tag{9}$$

Once $\hat{\boldsymbol{\beta}}$ is obtained, the OLS estimate of $\theta_i(U_t)$ is given by

$$\hat{\theta}_i^\star(U_t) = \sum_{j=1}^{M+K} \hat{\beta}_{ij} B_{j,K}(U_t),\ \text{for } i = 1, \ldots, p\,. \tag{10}$$

In particular for the FAR(p, d) model (2), the estimated model becomes

$$x_t = \sum_{i=1}^{p} \sum_{j=0}^{M+K} \hat{\beta}_{ij} B_{j,K}(x_{t-d}) + \epsilon_t\,. \tag{11}$$

where ϵ_t are residuals. Based on (11), we can construct multi-step-ahead forecasts for x_t. Let $\hat{x}_t(l)$ be the minimum mean square error prediction of x_{t+l}, then it can be carried out by iteratively implementing one-step-ahead prediction as

$$\begin{aligned}
\hat{x}_{t+1} &= \sum_{i=1}^{p} \left(\sum_{j=0}^{M+K} \hat{\beta}_{ij} B_{j,K}(\hat{x}_{t+1-d}) \right) \hat{x}_{t+1-i}\,, \\
&\vdots \\
\hat{x}_{t+l} &= \sum_{i=1}^{p} \left(\sum_{j=0}^{M+K} \hat{\beta}_{ij} B_{j,K}(\hat{x}_{t+l-d}) \right) \hat{x}_{t+l-i}\,,
\end{aligned} \tag{12}$$

where $\hat{x}_{t+l-d}$ equals x_{t+l-d} if $l - d \leq 0$.

2.3. Selection of Threshold Variable and Significant Variables

For ARMA(p, q) model, the impact of the past value x_{t-i}, $i = 1, \ldots, p$ on the current value x_t is direct and linear. However, in the real world, the impact may be related to another past value x_{t-d}. Compared to the ARMA(p, q) model, the FAR(p, d) model (2) is more flexible, where the impact of x_{t-i} on x_t can be related to the threshold variable x_{t-d}. The flexibility of the FAR(p, d) model can alleviate the model mis-specification problem and reduce forecasting error caused by choosing wrong models. In addition, the FAR(p, d) model does not involve the selection of exogenous variables. Due to these advantages, the FAR(p, d) model is widely used in empirical studies. Chen and Tsay (1993) [17] used the FAR(p, d) model to fit the monthly records of cases of chickenpox in New York City and the Wolf's annual sunspot numbers data set. Harvill and Ray (2005) [25] applied the FAR(p, d) model to forecast the U.S.GNP and the unemployment rate. Recently, the FAR(p, d) model has been extended and used in many fields, including survival analysis and risk management. For example, Xie et al. (2014) [26] proposed the varying-coefficient expectile (VCE) model to estimate the value at risk. In particular, by using the closing bid prices of the Euro in terms of the U.S. dollar, they applied the VCE model to fit the expectile of the weekly exchange return.

In this paper, we use the FAR(p, d) model to fit and forecast the inflation rate of China, and in the following we discuss some issues arising in the implementation of the FAR(p, d) model. In modelling and estimating the FAR(p, d) model, a key issue is choosing an appropriate threshold variable x_{t-d} and a set of significant variables $\{x_{t-i}\}_{i=1}^{p}$. This is equivalent to selecting the threshold lag d and the significant lag p with $d \leq p$. In practice, this can be achieved by a two-stage procedure. At the first stage, the significant lag p can be chosen based on subjective determination or by objective data driven methods such as information criteria. At the second stage, based on the chosen p, one decides d by minimizing the information criterion such as AIC (Akaike, 1974 [27]) or BIC (Schwarz, 1978 [28]) when p is fixed.

2.4. Selection of B-spline Basis Related Quantities

A crucial step in the B-spline estimation is to determine the B-spline basis, which is equivalent to determining the degree K, the number of knots M and the knots locations. The most commonly used degrees of the B-splines are 2 and 3, corresponding to the quadratic spline and the cubic spline. The determination of K can be conducted by the method of information criterion. Essentially, M is a smoothing parameter. As M increases, the spline function becomes less smooth leading to a more complicated model, while as M decreases, the spline function become more smooth with worse model fitting. In practice, M is chosen to balance the smoothness and model fitness. Huang and Shen (2004) [22] showed that the AIC outperforms other criterions such as the BIC and the modified cross-validation (Cai et al., 2000 [20]) in choosing M. Therefore in this paper we use the AIC to determine M.

For a given M, there are two popular ways to arrange the interior knots: $a < \xi_1 < \cdots < \xi_M < b$. One method gives the equally spaced knots, meaning that the distance between two adjacent knots is the same. The other one is the quantile knots, meaning that the knots locate at the $i/(M+1)$ sample quantiles ($i = 1, \ldots, M$) of the threshold variable. If the distribution of the threshold variable is not flat, the quantile knots are preferable (Huang and Shen, 2004 [22]).

3. Simulation Study

In this section, we check the fitting performance of the FAR model for different DGPs. Intuitively, if the underlying DGP follows a constant coefficient process, the linear ARMA will provide a better fit than the non-linear FAR model. On the other hand, if the underlying DGP follows time-varying coefficient process, then the FAR model can outperform than the ARMA model. Specifically, we compare the performance of the FAR and autoregressive (AR) model for different DGPs. The first case is a constant AR(2) model:

$$\text{Case I:} \quad \begin{cases} x_t = \theta_1^0 x_{t-1} + \theta_2^0 x_{t-2} + \varepsilon_t\,, \\ \theta_1^0 = 0.4, \quad \theta_2^0 = 0.5\,, \end{cases} \tag{13}$$

with $\varepsilon_t \overset{i.i.d.}{\sim} N(0, 1)$. The second case is the exponential AR model in Haggan and Ozaki (1981) [15] and Huang and Shen (2004) [22]:

$$\text{Case II:} \quad \begin{cases} x_t = \theta_1^0(x_{t-1})x_{t-1} + \theta_2^0(x_{t-1})x_{t-2} + \varepsilon_t\,, \\ \theta_1^0(x_{t-1}) = 0.138 + (0.316 + 0.982x_{t-1})e^{-3.89x_{t-1}^2}\,, \\ \theta_2^0(x_{t-1}) = -0.437 - (0.659 + 1.26x_{t-1})e^{-3.89x_{t-1}^2}\,. \end{cases} \tag{14}$$

In each simulation, a series of length T = 100, 400, 1000 and 2000 are drawn and the experiment is replicated 1000 times. For each case, we fit the generated data by both FAR(2,1) and AR(2) models. The coefficients of FAR(2,1) model and AR(2) model are estimated by the B-spline method and the OLS method, respectively. For each run s, the resulting estimates are denoted as $\hat{\theta}^{\star}_{i,s}(x_{t-d})$ and $\hat{\theta}_{i,s}$, respectively, $i = 1, 2$, $s = 1, \ldots, 1000$. All the simulations are run in R 3.4.3. To estimate the ARMA(p, q) model, we use the "arima" function built in the "stats" package.

To evaluate the fitting performance of both the FAR(2,1) model and the AR(2) model in the Cases I&II, we adopt the bias (BIAS) and the root mean square error (RMSE) of the corresponding estimates. Since the coefficients of the FAR model are functions of the threshold variable, the definitions of BIAS and RMSE of FAR models are different from those of AR models. Specifically, the BIAS and RMSE for Case I are defined as

$$\text{BIAS}_i = \begin{cases} \frac{1}{1000}\sum_{s=1}^{1000}(\hat{\theta}_{i,s} - \theta_i^0)\,, & \text{AR(2)}\,, \\ \frac{1}{1000}\sum_{s=1}^{1000}(\frac{1}{T-2}\sum_{t=3}^{T}\hat{\theta}^{\star}_{i,s}(x_{t-1}) - \theta_i^0)\,, & \text{FAR(2,1)}\,, \end{cases}$$

and

$$\text{RMSE}_i = \begin{cases} \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}(\hat{\theta}_{i,s} - \theta_i^0)^2}\,, & \text{AR(2)}\,, \\ \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}\frac{1}{T-2}\sum_{t=3}^{T}(\hat{\theta}^{\star}_{i,s}(x_{t-1}) - \theta_i^0)^2}\,, & \text{FAR(2,1)}\,, \end{cases}$$

respectively, $i = 1, 2$. The BIAS and RMSE for case II are defined as

$$\text{BIAS}_i = \begin{cases} \frac{1}{1000}\sum_{s=1}^{1000}(\frac{1}{T-2}\sum_{t=3}^{T}\hat{\theta}_{i,s} - \theta_i^0(x_{t-1}))\,, & \text{AR(2)}\,, \\ \frac{1}{1000}\sum_{s=1}^{1000}(\frac{1}{T-2}\sum_{t=3}^{T}\hat{\theta}^{\star}_{i,s}(x_{t-1}) - \theta_i^0(x_{t-1}))\,, & \text{FAR(2,1)}\,, \end{cases}$$

and

$$\text{RMSE}_i = \begin{cases} \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}\frac{1}{T-2}\sum_{t=3}^{T}(\hat{\theta}^{\star}_{i,s} - \theta_i^0(x_{t-1}))^2}\,, & \text{AR(2)}\,, \\ \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}\frac{1}{T-2}\sum_{t=3}^{T}(\hat{\theta}^{\star}_{i,s}(x_{t-1}) - \theta_i^0(x_{t-1}))^2}\,, & \text{FAR(2,1)}\,, \end{cases}$$

respectively, $i = 1, 2$. The results of BIAS and RMSE are reported in Tables 2 and 3.

Table 2. BIAS and RMSE of AR and FAR modelds for Case I.

T	Index	θ_1		θ_2	
		AR(2)	**FAR(2,1)**	**AR(2)**	**FAR(2,1)**
100	BIAS	−0.0332	−0.0486	−0.0454	−0.0051
400	BIAS	−0.0063	−0.0094	−0.0117	0.0042
1000	BIAS	−0.0025	−0.0036	−0.0053	0.0028
2000	BIAS	−0.001	−0.0015	−0.0026	0.0015
100	RMSE	0.0971	2.2412	0.1058	1.8270
400	RMSE	0.0449	1.8076	0.0455	1.5866
1000	RMSE	0.0278	1.7246	0.0281	1.5442
2000	RMSE	0.0191	1.6309	0.0194	1.5191

Table 3. BIAS and RMSE of the two models for case II.

T	Index	θ_1		θ_2	
		AR(2)	**FAR(2,1)**	**AR(2)**	**FAR(2,1)**
100	BIAS	0.0735	0.0003	0.0050	−0.0001
400	BIAS	0.0628	0.0001	0.0150	−0.0001
1000	BIAS	0.0609	0.0001	0.0177	0
2000	BIAS	0.0622	0.0001	0.0161	0
100	RMSE	1.8801	0.0038	2.8682	0.0060
400	RMSE	3.4010	0.0038	5.4510	0.0057
1000	RMSE	5.2670	0.0053	8.5164	0.0096
2000	RMSE	7.4267	0.0074	11.9953	0.0145

The results in Tables 2 and 3 show that, in terms of bias, when the DGP follows Case I, the performances for the FAR(2,1) and the AR(2) are similar, while the performance for FAR(2,1) is better than the AR(2) when the DGP follows Case II. Note that the RMSE for the FAR(2,1) specification in Case I is relatively large. This fact is reasonable because the squared sum of the difference between the constant true coefficients and the fitted time-varying coefficients becomes large as the time varies. Similarly, the RMSE for the AR(2) fitting for Case II are larger than that for the FAR(2,1). In terms of the RMSE, the difference between the performance of FAR(2,1) and AR(2) in Case II are larger than that in Case I. This result is due to the complicated time-varying coefficients of Case II. This fact implies that the FAR model works well when the underlying DGP is a constant coefficient, but the AR model works poor when the DGP is a time-varying coefficient model. In conclusion, the FAR model can reduce the risk of model mis-specification and thus can be widely used in empirical study when one has no prior information on the true model.

4. Empirical Analysis

4.1. Data Preprocessing

In this paper, inflation is measured in terms of the Consumer Price Index (CPI). The growth rate of CPI can be regarded as a proxy for the inflation rate. In particular, we use the monthly CPI data of China from Jan. 1995 to December 2017, with a total of 276 observations denoted as $\{x_t\}_{t=1}^{276}$. The data is displayed in Figure 1, showing that the CPI is quite large in the beginning of this period and drops down slowly. The inflation rate has been relatively stable at a level around 2% since 2012, which indicates that the development of the economy of China has stepped into the stage of "New Normal". Intuitively, the underlying inflation process may change since the beginning of the "New Normal" period.

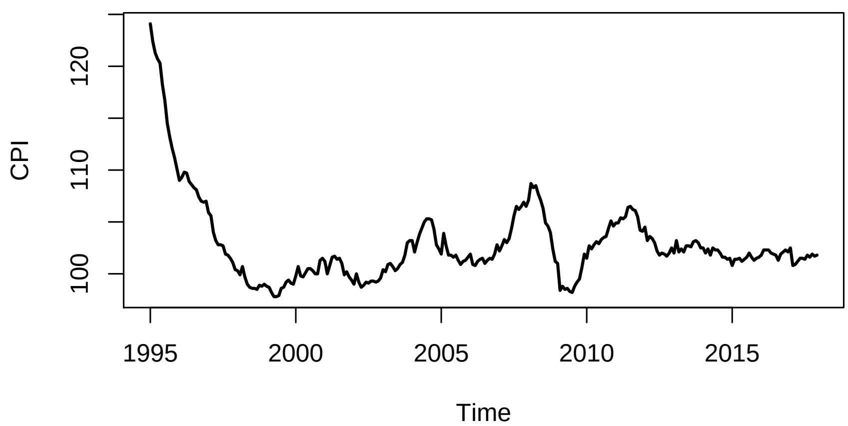

Figure 1. Monthly CPI data of China from January 1995 to December 2017.

In practice, we take the logarithm of the raw CPI data, and check the stationarity of the log-CPI data, at significance level 0.05. The result of the Phillips-Perron unit root test shows that the log-CPI data is nonstationary (with p-value 0.1494), but the first order difference of the log-CPI is stationary (with p-value 0). Therefore, the following analysis is based on the first order difference of the log-CPI denoted as

$$y_t = \log(x_t) - \log(x_{t-1}) \, . \tag{15}$$

The plot of y_t is shown in Figure 2. Parameter instability is also observed in our analysis of the data. We build an AR(1) model $y_t = \theta y_{t-1} + \epsilon_t$ and estimate the AR coefficient θ on an expanding window basis and rolling window basis with a 60 window-width. These estimates are plotted in Figure 3. It can be seen that the estimates of θ are quite variable. In conclusion, a non-linear FAR(p, d) model with time-varying coefficients is more reasonable and flexible than the linear ARMA model.

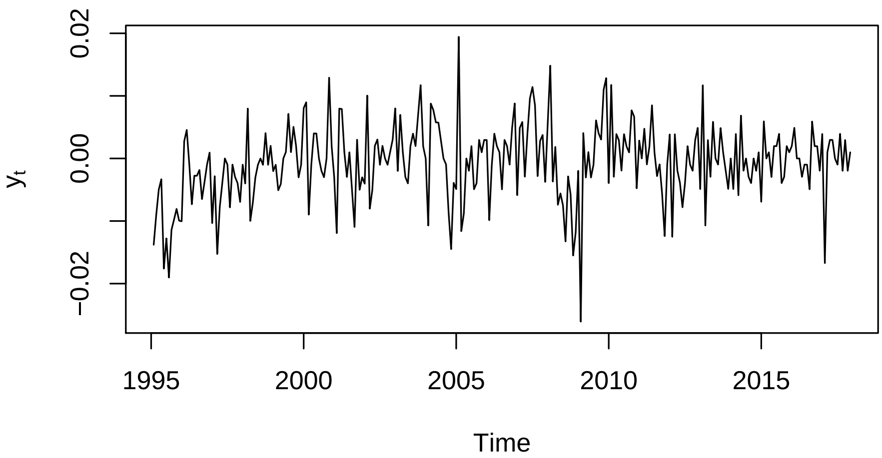

Figure 2. The first order difference of the monthly log-CPI.

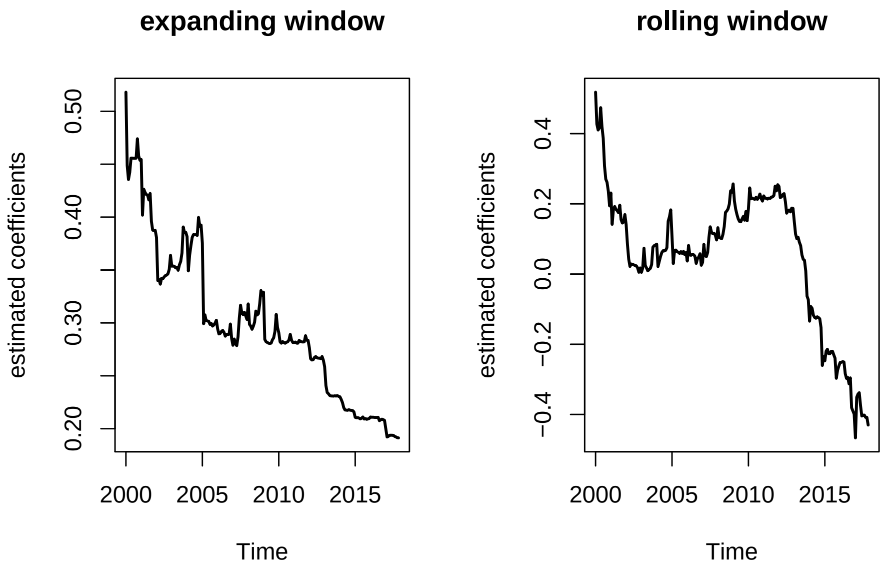

Figure 3. Estimates of θ. **Left:** expanding window; **right:** rolling window.

4.2. In-Sample Fit Analysis

We first use the FAR(p, d) model (2) to fit y_t given in (15). Based on the AIC values, we have $p = 4$. Then $d \leq p = 4$. For a given threshold lag d, the FAR$(4, d)$ model is estimated by the B-splines method. To construct the B-spline basis, the degree of the B-spline basis, the number and locations of the knots need to be determined. We consider different choices of the degree K and the number of interior knots M, i.e., $K = 2, 3$ and $M = 1, 2, 3, 4, 5$. For different values of (d, K, M), the locations of the knots ξ_i, $i = 1, \ldots, M$ are set to be the $i/(M+1)$ sample quantiles ($i = 1, \ldots, M$) of x_{t-d}. The B-spline basis can be calculated according to (5) and (6), and is then used to estimate the FAR(4, d) model by (9). We use the AIC criterion to determine the value of (d, K, M).

Table 4 reports the AIC values for different combinations of (d, K, M), showing that $(d, K, M) = (2, 2, 1)$ leads to the smallest value of AIC. Therefore, we use y_{t-2} as the threshold variable. $K = 2$ and $M = 1$ implies that the computational complexity is not large. The only internal knot is the median of $\{y_{t-2}\}_{t=3}^{T}$, while the boundary knots are $\max_{3 \le t \le T}\{y_{t-2}\}$ and $\min_{3 \le t \le T}\{y_{t-2}\}$, respectively. Thus, the resulting augmented vector of knots is (−0.02608, −0.02608, −0.02608, −0.02608, −0.0009, 0.0194).

Table 4. AIC values of FAR models with different (d, K, M).

AIC		$d = 1$	$d = 2$	$d = 3$	$d = 4$
$K = 2$	$M = 1$	−2002.20	−2008.77	−2006.09	−2001.36
	$M = 2$	−2002.38	−2002.27	−2003.53	−1994.18
	$M = 3$	−2005.80	−1996.11	−1997.89	−1989.72
	$M = 4$	−2002.97	−1989.58	−1991.41	−1993.06
	$M = 5$	−2007.87	−1985.52	−1983.82	−1996.12
$K = 3$	$M = 1$	−2006.04	−2003.56	−2003.16	−1994.44
	$M = 2$	−2006.58	−1996.92	−1996.13	−1995.36
	$M = 3$	−2002.43	−1990.42	−1989.67	−1990.75
	$M = 4$	−2006.26	−1987.78	−1985.55	−1985.35
	$M = 5$	−2002.85	−1982.75	−1983.23	−1987.87

The estimated FAR(4,2) model is

$$\hat{y}_t = -0.0003 + \hat{\theta}_1^{\star}(y_{t-2})y_{t-1} + \hat{\theta}_2^{\star}(y_{t-2})y_{t-2} + \hat{\theta}_3^{\star}(y_{t-2})y_{t-3} + \hat{\theta}_4^{\star}(y_{t-2})y_{t-4}, \tag{16}$$

where the estimated coefficients are $\hat{\theta}_i^{\star}(y_{t-2}) = \sum_{j=0}^{3} \hat{\beta}_{ij} B_{j,2}(y_{t-2})$, where $\hat{\beta}_{ij}$ are given in Table 5.

Based on the B-spline estimation results in Table 5, we plot the estimated functional coefficients of the FAR(4,2) model (16) in Figure 4. The fitted first order differenced monthly log-CPI $\hat{y}_t$ can be recursively obtained by (16). The Ljung-Box test shows that the residuals $\{\hat{e}_t = y_t - \hat{y}_t\}_{t=5}^{T}$ is a white noise sequence, indicating that the FAR(4,2) model provides a satisfactory model fitting. By substituting $\hat{y}_t$ into (15), we obtain the fitted CPI $\hat{x}_t$.

We also fit y_t by an ARMA(p, q) model, where $p, q \le 6$. The order p and q are determined by the AIC criterion, which gives $p = q = 5$. The estimated model is

$$\begin{aligned}\hat{y}_t =& -0.0008 + 0.6556y_{t-1} + 0.9923y_{t-2} - 0.7343y_{t-3} - 0.6319y_{t-4} + 0.5494y_{t-5} \\ & - 0.5173\epsilon_{t-1} - 0.1533\epsilon_{t-2} + 0.6032\epsilon_{t-3} + 0.9862\epsilon_{t-4} - 0.5607\epsilon_{t-5}.\end{aligned} \tag{17}$$

where $\epsilon_{t-1}, \ldots, \epsilon_{t-4}$ are residuals. Also, the residuals of model (17) are white noise, implying a satisfactory model fitting.

Table 5. B-spline estimation results of the FAR(4,2) model.

$\hat{\beta}_{ij}$	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$i = 1$	1.3807	−0.2424	0.3918	−1.4556
$i = 2$	0.9471	−0.2776	0.5190	−0.7262
$i = 3$	0.7882	0.2199	−0.2503	1.4473
$i = 4$	−1.6952	0.9303	−0.5024	1.1203

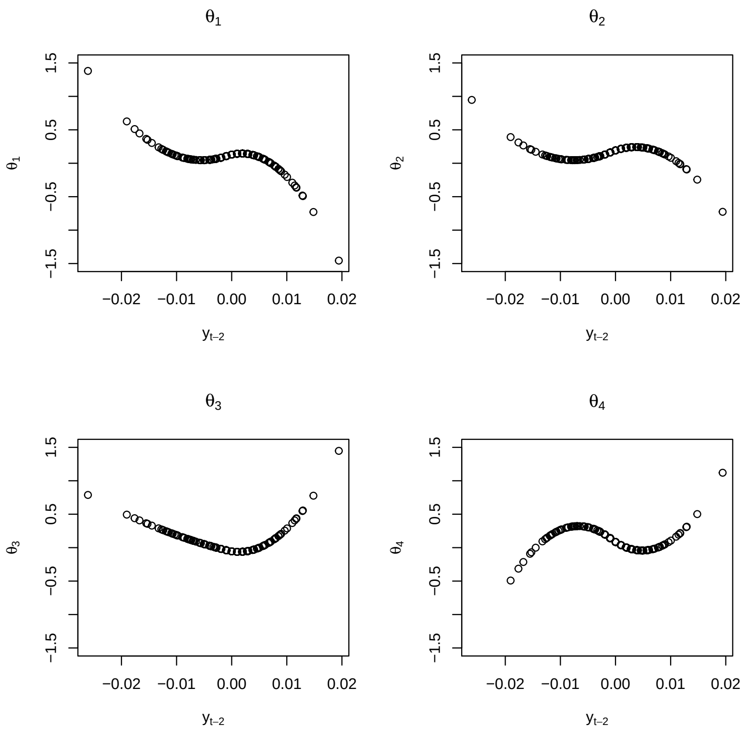

Figure 4. Estimation results of the coefficient functions.

Figures 5 and 6 display $\hat{y}_t$ and $\hat{x}_t$ given by the estimated FAR(4,2) and ARMA(5,5) models, respectively. From Figure 6, both the FAR(4,2) and the ARMA(5,5) describe the main characteristic of data quite well. In particular, they capture the three falling and rising processes of the inflation since 1995. Intuitively, the fitted ARMA(5,5) is more fluctuated, meaning that the model overestimates the peak and underestimates the trough, while the fitted FAR(4,2) model is more smoothing, i.e., the FAR(4,2) model underestimates the peak and overestimates the trough. Thus, it is expected that the ARMA(5,5) model fit the CPI data better during the fluctuating period, while the FAR(4,2) model performs better during the stable period. From Figure 6, the inflation rate during 1995 to 2011 fluctuates heavily, and the fitted CPI by using the ARMA(5,5) model is closer to the real CPI than that given by the FAR(4,2) model. However, for the period from 2012 to 2017, the inflation rate is quite stable around the level of 2%, and the FAR(4,2) model performs better.

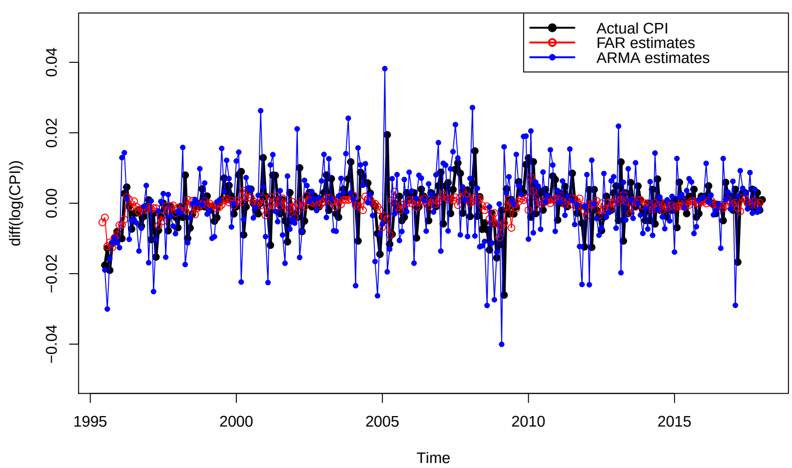

Figure 5. Estimated first order difference of the log-CPI given by the FAR(4,2) model and the ARMA(5,5) model.

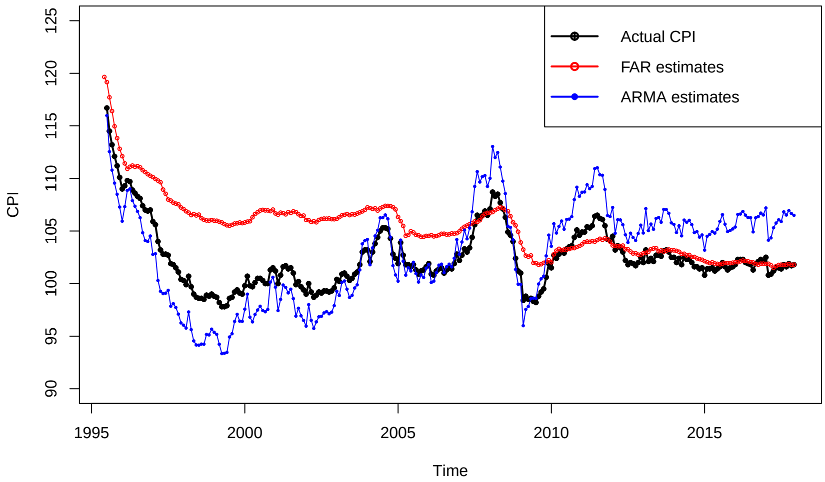

Figure 6. Estimated CPI given by the FAR(4,2) model and the ARMA(5,5) model.

We compare the fitting performance of the FAR(4,2) model and the ARMA(5,5) model for x_t by using mean absolute errors (MAE) and RMSE defined as

$$\mathrm{MAE} = \frac{1}{T}\sum_{i=1}^{T} |x_t - \hat{x}_t|, \tag{18}$$

and

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{T}(x_t - \hat{x}_t)^2}{T}}. \tag{19}$$

The definitions of MAE and RMSE for y_t are similar and thus are omitted. Furthermore, to compare predictive accuracy, we employ the Diebold and Mariano test (denoted as DM test hereafter) proposed in Diebold and Mariano (1995) [29], based on the corresponding MAE and RMSE. We only introduce the DM test based on the MAE for simplicity. Define the forecasting error of x_t as $e_t = x_t - \hat{x}_t$ and the loss function $L(e_t) = |e_t|$. Based on $L(e_t)$, the forecasting loss difference between ARMA and FAR is $d_t = L(e_t^A) - L(e_t^F)$. Here, the null hypothesis is that the ARMA model has equal predictive accuracy as the FAR model, which is equivalent to $\mathrm{E}(d_t) = 0$. Let $\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$, then the DM statistic is given by

$$DM = \frac{\bar{d}}{\hat{\sigma}_{\bar{d}}},$$

where $\hat{\sigma}_{\bar{d}}$ is a consistent estimate of the standard deviation of $\bar{d}$. The alternative can be set as the forecast of FAR is more accurate than that of ARMA. The procedure proceeds as follows. First, we calculate DM test statistic and the two-sided p-value based on the limiting standard normal distribution. Then if the value of the test statistic is positive (negative), the one-sided p-value is just one half of the two-sided p-value (one minus one half of the two-sided p-value). Similarly, if the alternative is that the forecast of ARMA is more accurate than that of FAR, and the value of the test statistic is positive (negative), the one-sided p-value is one minus a half of the two-sided p-value (one half of the two-sided p-value).

In addition, we divide the in-sample forecasts into two halves: the first half ranges from January 1995 to September 2006 and the second half ranges from October 2006 to December 2017. The two-sided testing results are given in Table 6. We first discuss the forecast of y_t. With the full sample, the MAE and RMSE of ARMA(5,5) model are almost the same as those of FAR(4,2) model for fitting y_t. The scenario is similar for the first period data and the second period data. This fact is in line with the DM test results, which shows that in most cases, the fitting accuracies of the two models are similar under the 5% significance level.

Then we consider the results for x_t. For the full sample, the MAE of ARMA(5,5) is 15.8% smaller than those of FAR(4,2) model for fitting x_t. This improvement increases to 58.19% with the first period data, which can be also verified by the larger absolute values of the DM statistic. The DM test result implies that the ARMA(5,5) model provides better fitting accuracies than the FAR(4,2) model. However, for the more stable second period data, the MAE of ARMA(5,5) is three times of that of FAR(4,2) model. This superiority can be also verified by the DM test. Therefore, the FAR(4,2) model has better fitting accuracy for more stable data. Similar conclusions can be obtained based on RMSE.

Combining the previous analysis, it is expected that the FAR model will provide more accurate prediction for x_t when China are stepping into the "New Normal" stage.

Table 6. Comparion of MAE and RMSE in ARMA(5,5) and FAR(4,2) model.

		$\hat{y}_t$				$\hat{x}_t$			
		FAR	**ARMA**	**DM Test**	***p*-Value**	**FAR**	**ARMA**	**DM Test**	***p*-Value**
Full sample	MAE	0.0039	0.0041	−**1.2065**	0.2287	3.0783	2.5918	−**2.6366**	0.0083
	RMSE	0.0052	0.0055	−**1.9427**	0.0531	3.9805	2.9115	−**6.0828**	0.0000
First half	MAE	0.0041	0.0041	−**0.3154**	0.7527	5.1302	2.1452	−**38.4356**	0.0000
	RMSE	0.0055	0.0052	−**1.4703**	0.1427	5.4395	2.5039	−**27.1525**	0.0000
Second half	MAE	0.0042	0.0039	−**1.4839**	0.1402	1.0337	3.0451	**11.3261**	0.0000
	RMSE	0.0056	0.0051	−**1.6909**	0.0932	1.4738	3.2722	**10.6661**	0.0000

"Full sample" means the loss function and the DM test are calculated based on the full sample period ranging from June 1995 to December 2017. "First half" and "Second half" means the results are calculated by using the first half samples and the second half samples, respectively. The calculated DM statistic (in bold) and the corresponding two-sided p-value are reported.

4.3. Out of Sample Forecast

From the in-sample analysis, the predictive performance of each model is not the same for periods before and in the stage of "New Normal". Compared to the in-sample fitting, the out-of-sample forecasting of a model is more important, since precise forecasts of the inflation rate are crucial for economic agents (e.g., investors, consumers) as well as for economic policy decision makers. In particular, we are interested in the inflation rate for the period of "New Normal". Thus, we divide the whole sample into two parts: the first part covers data from January 1995 to October 2013, and the second part covers data from November 2013 to December 2017. We use the rolling window method to obtain the future CPI value. That is, when the forecast proceeds, the estimation window rolls forward by adding one new data and dropping the most distant data. In this way, the size of the estimation window remains the same. The one-step-ahead forecast is implemented based on the estimated FAR(4,2) model (16) and the ARMA(5,5) model (17). To get an h-step-ahead forecast ($h > 1$), we can iteratively implement one-step-ahead forecast h times as given in (12). In the following, we consider different values of forecast horizons $h \in \{1, 3, 6, 9, 12, 15, 18, 21, 24\}$. The forecasting results are shown in Figure 7.

The figure shows that the forecasts based on the FAR(4,2) model are more smoothing, while the forecasts based on the ARMA(5,5) model are more fluctuated. Since the true CPI data in the forecasting period are stable, the FAR(4,2) model provide more accurate forecasts. It is also shown that as h increases, the forecasts for both models become more fluctuated and less accurate.

To evaluate the forecasting accuracy, we use the forecasting MAE and RMSE defined as

$$\text{MAE}^h = \frac{1}{N} \sum_{t=1}^{N} |x_t - \hat{x}_t^h| , \tag{20}$$

and

$$\text{RMSE}^h = \sqrt{\frac{\sum_{t=1}^{N}(x_t - \hat{x}_t^h)^2}{N}}, \tag{21}$$

respectively, where x_t is the true CPI, and $\hat{x}_t^h$ is the h-step-ahead forecast of CPI, N is the total number of forecasts. We also conduct the DM test for comparison. In Table 7, we report the MAE, the RMSE, the DM statistic and the p-value with alternative hypothesis that the two models have different predictive accuracies. The forecast horizon is fixed at $h \in \{1, 3, 6, 9, 12, 15, 18, 21\}$. It can be observed that at short horizon levels (i.e., $h \in \{1, 3\}$), the MAE and RMSE of the two models are comparable. When the horizon level becomes large (i.e., $h \in \{6, 9, 12, 15, 18, 21\}$), the MAE and RMSE of FAR(4,2) model are lower than those of ARMA(5,5) model. Moreover, as the horizon level h increases, the improvement of FAR(4,2) model becomes larger. This phenomenon is also detected by the DM test. Specifically, at 5% significance level, the FAR(4,2) model outperforms the ARMA(5,5) model for $h \in \{6, 9, 12, 15, 18, 21\}$, while the predictive accuracies of the two models are similar for $h \in \{1, 3\}$. This fact implies that the FAR model is better for moderate and long-term inflation rate forecasting and is comparable to ARMA model for short-term inflation rate forecasting.

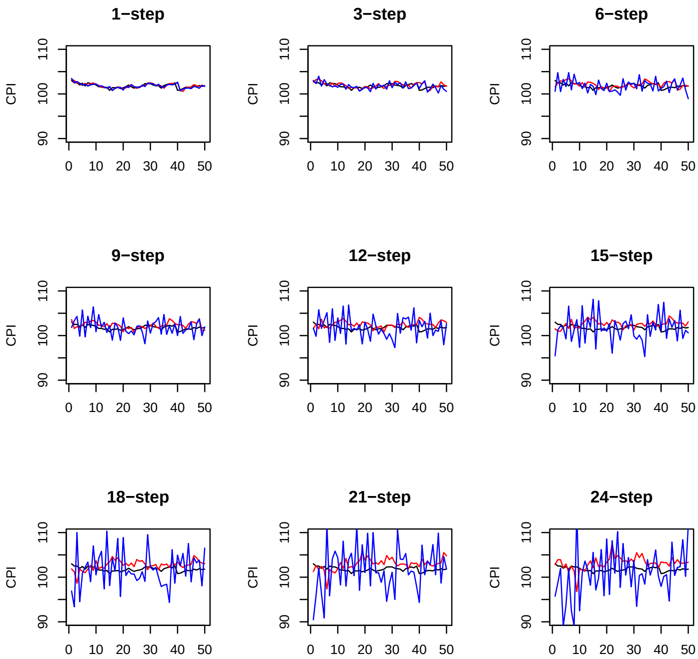

Figure 7. Multi-step-ahead forecasts of the monthly CPI data given by the FAR model and the ARMA model (black line: true CPI; blue line: ARMA; red line: FAR).

Table 7. MAE and RMSE of multi-step-ahead forecasts.

	h	1	3	6	9	12	15	18	21
MAE	ARMA(5,5)	0.3057	0.6052	1.1990	1.4782	2.1926	2.5480	3.3967	4.2165
	FAR(4,2)	0.2969	0.5278	0.6196	0.7560	0.9500	1.1295	1.3590	1.5832
	DM test	**0.3244**	**1.3658**	**4.9881**	**2.9283**	**4.9559**	**4.8971**	**3.8484**	**2.7157**
	p-value	0.7470	0.1782	0.0000	0.0052	0.0000	0.0000	0.0003	0.0091
MSE	ARMA(5,5)	0.4246	0.7447	1.4027	1.7951	2.5945	3.2504	4.2081	5.2963
	FAR(4,2)	0.4102	0.6656	0.8286	0.8883	1.2054	1.3406	1.6212	1.8738
	DM test	**0.6118**	**1.2482**	**3.4193**	**3.4778**	**5.3781**	**5.4008**	**2.7253**	**2.2040**
	p-value	0.5435	0.2179	0.0013	0.0011	0.0000	0.0000	0.0089	0.0323

The calculated DM statistic (in bold) and the corresponding two-sided p-value are reported.

5. Conclusions

The inflation rate is a critical quantity for both policy-makers and economic researchers. Hence, forecasting the inflation rate has long attracted the interests from various fields. In this paper, we apply the FAR model to forecast the inflation rate after the economy of China stepping into a new stage of "New Normal". The FAR model belongs to the semi-parametric non-linear time series model, which has three main advantages. First, compared to the traditional linear time series model, this model can describe the non-linear dynamics of the underlying process, which is common for many real time series data. Second, compared to the fully parametric non-linear time series model for example the threshold autoregressive model, the FAR model is more flexible and avoids the problem of model mis-specification. Third, the FAR model attains a satisfactory forecasting power, particularly for data without much fluctuations. The last advantage matches the characteristic of the inflation rate of China during the "New Normal" period. To estimate the functional coefficients of the FAR model, we adopt the B-spline method, which is numerically stable and can be obtained recursively. Thus, this estimation method largely reduces the computational burden and can be applied to handle huge amount of data.

Code for this illustration was written in R (R Development Core Team, 2010) and is available upon request from the authors.

Author Contributions: Q.L. and C.C. designed and performed the simualtion study and analyzed the data; M.W. and K.C. designed the empirical study, contributed analysis tools and wrote the paper.

Acknowledgments: We thank the editor and three anonymous referees for their helpful comments and useful references, which led to an improved version of this paper. Man Wang's work was supported in part by Social Science Foundation of Shanghai (No. 2017EJB009), the Fundamental Research Funds for the Central Universities (No. 18D110801) and the Donghua University Special Research Funds for Social Science (No. 108-10-0108117).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bampinas, G.; Panagiotidis, T. Hedging inflation with individual US stocks: A long-run portfolio analysis. *N. Am. J. Econ. Financ.* **2016**, *37*, 374–392. [CrossRef]
2. Stock, J.H.; Watson, M.W. Forecasting inflation. *J. Monet. Econ.* **1999**, *44*, 293–335. [CrossRef]
3. Atkeson, A.; Ohanian, L.E. Are Phillips Curve useful for forecasting inflation? *Fed. Reserv. Bank Minneap. Q. Rev.* **2001**, *25*, 2–11.
4. Mcadam, P.; Mcnelis, P. Forecasting inflation with thick models and neural networks. *Econ. Model.* **2005**, *22*, 846–867. [CrossRef]
5. Matheson, T.D. Phillips curve forecasting in a small open economy. *Econ. Lett.* **2006**, *98*, 161–166. [CrossRef]
6. Sekine, T. *Modeling and Forecasting Inflation in Japan*; IMF Working Paper No.01/82; International Monetary Fund: Washington, DC, USA, 2001.
7. Ramakrishman, U.; Vamvakidis. A. *Forecasting Inflation in Indonesia*; IMF Working Paper No.02/111; International Monetary Fund: Washington, DC, USA, 2002.
8. Ang, A.; Bekaert, G.; Wei, M. Do macro variables, asset markets, or surveys forecast inflation better? *J. Monet. Econ.* **2007**, *54*, 1163–1212. [CrossRef]
9. Lekkos, I.; Milas, C.; Panagiotidis, T. Forecasting interest rate swap spreads using domestic and international risk factors: Evidence from linear and non-linear models. *J. Forecast.* **2007**,*26*, 601–619. [CrossRef]
10. Milas, C.; Otero, J.; Panagiotidis, T. Forecasting the spot prices of various coffee types using linear and non-linear error correction models. *Int. J. Financ. Econ.* **2004**, *9*, 277–288. [CrossRef]
11. Bos, C.S.; Franses, P.H.; Ooms, M. Inflation forecast intervals and long memory regressions models. *Int. J. Forcast.* **2000**, *18*, 243–264. [CrossRef]
12. Tong, H. *Nonlinear Time Series: A Dynamical System Approach*; Oxford University Press: Oxford, UK, 1990.
13. Chen, S.X.; Lei, L.; T,Y. Functional coefficient moving average model with applications to forecasting chinese CPI. *Stat. Sin.* **2016**, *26*. 1649–1672. [CrossRef]
14. Granger, C.W.; Andersen, A. On the invertibility of time series models. *Stoch. Proc. Appl.* **1978**, *8*, 87–92. [CrossRef]

15. Haggan, V.; Ozaki, T. Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika* **1981**, *68*, 189–196. [CrossRef]
16. Fan, J.Q.; Yao, Q.W. *Nonlinear Time Series: Nonparametric and Parametric Methods*; Springer: New York, NY, USA, 2003.
17. Chen, R.; Tsay, R.S. Functional coefficient autoregressive model. *J. Am. Stat. Assoc.* **1993**, *88*, 298–308.
18. Hastie, T.; Tibshirani, R. Varying-coefficient models. *J. R. Stat. Soc. B* **1993**, *55*, 757–796.
19. Chen, B.; Hong, Y. Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica* **2012**, *80*, 1157–1183.
20. Cai, Z.W.; Fan, J.Q.; Yao, Q.W. Functional-coefficient regression models for nonlinear time series. *J. Am. Stat. Assoc.* **2000**, *95*, 941–956. [CrossRef]
21. Chen, R.; Liu, L.M. Functional coefficient autoregressive models: Estimation and tests of hypotheses. *J. Time Ser. Anal.* **2001**, *22*, 151–173. [CrossRef]
22. Huang, J.Z.; Shen, H.P. Functional coefficient regression models for non-linear time series: A polynomial spline approach. *Scand. J. Stat.* **2004**, *31*, 515–534. [CrossRef]
23. De Boor, C. On Calculating with B-Spline. *J. Approx. Theory* **1972**, *6*, 50–62. [CrossRef]
24. Schumaker, L. *Spline Functions: Basic Theory*; Cambridge University Press: Cambridge, UK, 2007.
25. Harvill, J.L.; Ray, B.K. A note on multi-step forecasting with functional coefficient autoregressive models. *Int. J. Forcast.* **2005**, *21*, 717–727. [CrossRef]
26. Xie, S.; Zhou, Y.; Wan, A.T.K. A Varying-Coefficient Expectile Model for Estimating Value at Risk. *J. Bus. Econ. Stat.* **2014**, *32*, 576–592. [CrossRef]
27. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
28. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
29. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263.