Forecasting Distribution of Inflation Rates: Functional Autoregressive Approach*

Kausik Chaudhuri
University of Leeds, UK

Minjoo Kim
University of Glasgow, UK

Yongcheol Shin
University of York, UK

Final version: 30 November 2014

Abstract

In line with the recent developments on the statistical analysis of functional data, we develop the semiparametric functional autoregressive (FAR) modeling approach to the density forecasting analysis of national inflation rates using sectoral inflation rates in the UK over the period January 1997-September 2013. The pseudo out-of-sample forecasting evaluation and test results provide an overall support to superior performance of our proposed models over the aggregate autoregressive models and their statistical validity. The fan-chart analysis and the probability event forecasting exercise provide a further support for our approach in a qualitative sense, revealing that the modified FAR models can provide a complementary tool for generating the density forecast of inflation, and analyse the performance of the central bank in achieving announced inflation target. As inflation targeting monetary policies are usually set with recourse to the medium-term forecasts, our proposed work may provide policymakers with an invaluably enriched information set.

JEL Classification: C14, C53, E31.

KEYWORDS: Time-varying Cross-sectional Distribution, Functional Autoregression, Nonparametric Bootstrap, Density and Probability Forecasting of the UK Inflation.

*We are mostly grateful to the editor, Arnaud Chevalier, the associate editor, and two anonymous referees for their valuable comments and suggestions. We are also grateful to Murali Agastya, Chiyoung Choi, Jinho Choi, Matthew Greenwood-Nimmo, Chungrak Kim, Inmoo Kim, Soyoung Kim, Taehwan Kim, Kevin Lee, Myungjae Lee, Simon Lee, Guay Lim, Amy Mise, Viet Nguyen, Les Oxley, Joon Park, Yangsoo Park, Kevin Reilly, Andrew Scott, Ron Smith, Seungjoo Song, Sam Yoo, seminar participants at the Bank of Korea, Universities of Busan, Korea, Leeds, Leicester, Melbourne, Yonsei and York, and conference delegates at Far East and South Asia Meeting of the Econometric Society, University of Tokyo, August 2009 and KDI/KAEA Conference, Korea Development Institute, August 2009 for their helpful comments. Kim (second author) acknowledges that this paper is an extension of the second chapter of his PhD thesis submitted to the University of Leeds in 2011. The usual disclaimer applies. Shin (third author) gratefully acknowledges partial financial support from the ESRC (Grant No. RES-000-22-3161). The usual disclaimer applies.

1 Introduction

Monetary authorities around the world began to adopt inflation targeting as the monetary policy framework, by which the bank commits to keep inflation within a narrow pre-defined level, in the 1990s. In this policy framework, an inflation forecast is a key input in the decision making process of the Monetary Policy Committee (as it is known in the UK), since it signals that a potential change in policy may be required to ensure that the inflation rate does not move outside its target range. The importance of providing additional information on the uncertainty surrounding forecasts of key macroeconomic variables has been increasingly recognised, e.g., Giordani and Söderlind (2003). Point forecasts convey limited information and work well only in restrictive cases (Granger and Pesaran, 2000). More generally, knowledge about the precision of forecasts enables policymakers to motivate actions based on these forecasts, and helps in achieving a more balanced evaluation of both forecasts and policy in the public arena (Casillas and Bessler, 2006). Monetary policy decisions should explicitly accommodate the uncertainty surrounding point forecasts. The BOE produces a quarterly inflation report in which this uncertainty is conveyed using fan charts over a two-year horizon, with bands of various shades of red illustrating the range of probable outcomes (http://www.bankofengland.co.uk/publications/inflationreport/irfanch.htm). Similarly, the European Forecasting Network and the Fed provide such forecasts for a set of key macroeconomic variables. Although this current practice is clearly an important step toward acknowledging the significance of forecast uncertainties in the decision making process, it is difficult for independent researchers to replicate them due to the subjective manner in which the central bank accommodates forecast uncertainty. In this paper, using sectoral inflation data, we use the informational contents present in cross-sectional higher order moments and the time-variation of the cross-sectional density to develop an inflation forecasting framework.

The time-varying and non-normal distribution of inflation rates may have dramatic consequences for the optimal conduct of monetary policy. Accordingly, reliable forecasts have to be computed dynamically to account for changes in the distribution. Previous studies have used the autoregressive (AR) process to model the inflation rate or the autoregressive conditional heteroskedastic (GARCH) process (Engle, 1982; Bollerslev, 1986) for modeling inflation uncertainty. A number of studies have examined the role of time-varying higher-order moments in the analysis of financial and macroeconomic data (e.g. Hansen, 1994; Bryan et al., 1997; Harvey and Siddique, 2000; Roger, 2000; Jondeau and Rockinger, 2003). In the case of inflation, Bryan et al. (1997) document fat-tailed properties and Roger (2000) provides evidence of rightward skewness. It follows that analysis based on aggregate data may be severely biased given the presence of informational heterogeneity across the sectors as documented by Pesaran and Smith (1995) and Hsiao et al. (2005).

We develop a flexible framework for utilising both the informational content of the higher-order moments and the time-variation of the cross-sectional distribution of sectoral inflation rates. Specifically, we follow Bosq (2000), Cardot et al. (1999, 2007), Mas (2007) and Park and Qian (2007, 2012), and develop a semi-parametric functional autoregressive (hereafter, FAR) modeling strategy for forecasting the time-varying cross-sectional density. In particular, the cross-sectional densities are treated as time series of the functional data and their dynamic nature is estimated via an autoregressive model in the functional space. As the dependence of the variations over sectors is being modeled nonparametrically, we do not impose any assumptions on the structure of the distributions, or on the number of dimensions in which the distributions may vary over time. We then employ the autoregressive operator to specify the time-dependence of the distribution function, thereby allowing all the moments to depend on all the past moments in a flexible manner.

Our approach is in line with the recent developments in the statistical analysis of functional data. Bowsher and Meeks (2008) and Kargin and Onatski (2008) employ these models for forecasting the yield curve. Notice, however, that these existing applications focus on time variation only. Despite the FAR model being regarded as an unusually powerful forecasting tool, we are unaware of any study that has applied this strategy explicitly to the high profile issue of low-frequency inflation forecasting. It is here that we contribute to the literature by developing a flexible semi-parametric FAR modeling for the construction of density forecasts of national inflation rate on the basis of sectoral inflation data which varies across cross-section as well as over time periods.

Attempts have also been made to utilise information available at the disaggregate level by using the data from the Survey of Professional Forecasters (SPF) (e.g., Zarnowitz and Llambros, 1987; Lahiri and Liu, 2006) and by using the CPI confidential microdata of Bureau of Labor Statistic (BLS) (e.g., Bils and Klenow, 2004; Klenow and Malin, 2010). The SPF data utilises inflation forecasts from different forecasters and draws the density of the inflation rate. The density of the SPF data is reported as a histogram and thus it is not straightforward to evaluate its higher order moments (Engelberg et al., 2009; Clements 2010). Two additional problems also arise when working with the SPF data: heterogeneity of forecasters and changes in the composition of the panel over time. We contribute to the existing literature by developing a flexible framework that can readily accommodate the unbalanced nature of such datasets.

Research using the BLS data has focused on examining the distribution of prices, but no attempt has yet been made to forecast the density of aggregate inflation. In principle, we can forecast the density function of the national inflation rate utilising information at either the sectoral or the firm level by combining the nonparametric kernel density estimation of sectoral inflation rates, the functional autoregressive modeling of cross-sectional densities, and the evaluation of density forecast of inflation rate via nonparametric bootstraps. Our hybrid approach is then expected to provide additional findings compared to the existing work with the BLS data.

Based on a detailed exploratory analysis of the time-varying cross-sectional moments of the UK sectoral CPI inflation rates over January 1997 and September 2013, we find that the first four moments are highly persistent and interrelated. The national inflation rate is highly correlated with both relative price variability (RPV) and relative price skewness (RPS), though the direction of the association is contingent on the location of the inflation relative to adaptive inflation expectations (defined as a learning process through which future expectations can be formed on the basis of what has happened in the past). The mean inflation is positively correlated with RPV and RPS when current inflation exceeds expected inflation and negatively correlated otherwise, especially in the pre-crisis period (1997-2007). These findings provide a partial support to Friedman (1977) and Cukierman and Meltzer (1986), and favor the menu cost explanation of Ball and Mankiw (1995) regarding the sluggish adjustment of prices in response to aggregate shocks only under the higher inflation regime. This implies that firms will change the price only if the profit from the relative price adjustment is larger than the menu cost (i.e., the cost that accrues as a result of changing their price). On the contrary the negative association observed in the low inflation regime supports the exogenous form of downward nominal rigidity advanced by Tobin (1972), in which case prices of most goods and services do not change immediately following a shock to the firms' relative prices. For the post-crisis period (2008-2013), we still observe the positive correlation between inflation and RPV under the higher inflation regime, and a negative one under the lower inflation regime. Interestingly, however, the post-crisis relationship between inflation rate and RPS becomes an inverse U-shape. But, the degree of associations tends to be substantially weaker in the post-crisis period. Overall, our finding of the asymmetric U-shaped pattern that characterises the relationship between inflation rate and RPV is generally consistent with recent studies by Chen et al. (2008), Choi (2010) and Chaudhuri et al. (2013).

Our data spans the recent period of the global financial and the European debt crisis during which inflation rates are found to be consistently above target at times. Therefore, to evaluate the relative performance of our proposed methodology, we repeat the forecasting exercise for different sample periods: the pre-crisis, the crisis and the post-crisis periods. We conduct a number of forecasting evaluation exercises by considering different functional models. These include three basic models, namely the functional autoregressive (FAR), the functional average (AVE) and the functional martingale (LAST) models. AVE selects the average of all the observed distributions and LAST selects the last distribution for forecasting purpose. Following Clements and Hendry (2002, 2006), we have considered modifications either by intercept shifting or by taking the first-difference in order to incorporate (possible) structural breaks and the high persistence of the moments of sectoral inflation rates. Thus, we introduce the nine additional modified models, namely, the 3-, 6-, 9- and 12-month moving averages of FAR and AVE models (denoted as FAR3M, FAR6M, FAR9M, FAR12M, AVE3M, AVE6M, AVE9M and AVE12M) and the differenced FAR model (DFAR). Furthermore, we have also proposed a robust bootstrap scheme as a complementary to the standard one in order to evaluate the empirical distributions in a more robust manner, especially in the presence of extreme events such as the global financial crisis. This is generally in line with the generalised extreme value distribution theory as introduced by Embrechts et al. (1997) and applied to financial applications by Longin (2000) and McNeil and Rudiger (2000). We find that our proposed modifications seem to handle the data for both the pre-crisis and the post-crisis periods reasonably well, therefore justifying their application from a practical point of view.

The evaluation results of pseudo out-of-sample forecasting (e.g., Elliot and Timmermann, 2008; Faust and Wright, 2012) reveal that the performances of unmodified models such as FAR and AVE are rather poor, and that they are sometimes dominated by the benchmark aggregate AR and ARCH-M models. We find that modified models, especially DFAR, FAR3M and AVE3M, are the best among the functional models, and they strongly dominate the benchmark models. Furthermore, we have evaluated the density forecasting performance using the probability integral transformations (PIT) proposed by Diebold et al. (1998), and accessed the accuracy of the interval forecasting by applying the conditional coverage test developed by Christoffersen (1998). Both test results provide an overall support to the statistical validity of our modified functional autoregressive (FAR) modeling approach using the disaggregate data. In particular, DFAR shows strong and robust statistical evidence and can forecast the density of the national inflation rates even during the crisis period.

The fan-chart analysis demonstrates that realised inflation rates are well within the confidence band with forecasted mean inflation tracking actual one reasonably well during both the pre-crisis and the post-crisis period. In particular, our fan-charts are remarkably similar to those published by the BOE quarterly bulletin in fourth quarter of 2004 and in November 2010. Furthermore, we examine the probability of achieving the inflation target of $\pi < 2\%$ and $1\% < \pi < 3\%$ as announced by BOE, finding that the probability event forecasting exercises can capture the actual violations almost accurately, especially during the post-crisis period. We also evaluate the probability forecasting performance by quadratic probability scores (QPS, Brier, 1950) and the calibration tests (Seillier-Moiseiwitsch and Dawid, 1993; Diebold and Lopez, 1996). Both results provide an overall support for modified functional models, especially the satisfactory performance of DFAR over the short forecasting horizons up to 6-8 months. In sum these analyses seem to support our modeling approach in a qualitative sense, revealing that our proposed approach is able to provide a complementary tool for generating the density forecast for inflation and analyse the performance of the central bank in achieving announced inflation target.

Our results show that utilisation of disaggregate data using a semi-parametric model provides more useful information in forecasting the aggregate inflation compared to an aggregate parametric model, which is likely to result in misleading forecast. As inflation targeting monetary policies are usually set with recourse to medium-term forecasts, our work on forecasting the time-varying distribution of sectoral inflation rates may provide policymakers with an invaluably enriched information set.

The rest of the paper is organised as follows. Section 2 introduces the FAR methodology. The exploratory data analysis of the sectoral inflation rates is reported in Section 3. The main estimation and forecasting results based on the FAR models are provided in Section 4. Section 5 concludes. Appendix provides the computation algorithms for the semiparametric FAR model estimation and the bootstrap-based inference in details.

2 Functional Autoregressive Distribution

The statistical foundation of the FAR model is comprehensively introduced by Bosq (2000). Cardot et al. (1999, 2007) and Mas (2007) refine the asymptotic theory. The early applications are in forecasting climate patterns such as temperatures (Besse et al., 2000) and ozone (Damon and Guillas, 2002; Aneiros-Perez et al., 2004). Recently, the application has been adopted in finance. Laukaitis (2008) utilises the FAR model to forecast the intraday cash flow and intensity of the transaction in a credit card payment system. Bowsher and Meeks (2008) and Kargin and Onatski (2008) apply it to forecast a yield-curve as the function of maturity. See also Park and Qian (2007, 2012) for forecasting the nonparametric density function of the intraday stock returns.

In line with these developments, we aim to develop the novel semiparametric FAR modeling strategy for constructing the density forecast of the national inflation rate using the cross-sectional sectoral inflation rates as follows: (i) The cross-sectional density of sectoral inflation rates is estimated non-parametrically at each time period by a weighted kernel density estimator; (ii) These time-varying cross-sectional densities are then modeled by FAR in a functional space; (iii) Finally, the empirical distribution of the forecasted national inflation rate is generated via the nonparametric bootstrap techniques. Notice that all previous empirical studies focus on the time variation only whereas our approach combines both cross-sectional and time variations.

Let (Ω, μ) be a measurable space where Ω denotes the set of all sectors in an economy and μ is the Lebesgue measure on Ω. We assume that the number of sectors in Ω is either finite or infinite, but it should be denumerable in practice (e.g. $\mathbb{D} \subseteq \mathbb{R}$). Denote $x_t(\omega)$ as the inflation rate at time t in a sector, $\omega \in \Omega$, and $f_t : \Omega \rightarrow \mathbb{R}^+$ as a density function, satisfying $\int_\Omega f_t(\omega)\, d\mu(\omega) = 1$, as it measures the relative weight assigned to each sector. Then, the mean inflation, denoted π_t, can be defined by

$$\pi_t = \int_\Omega \underbrace{x_t(\omega)}_{\text{inflation rate}} \underbrace{f_t(\omega)\, d\mu(\omega)}_{\text{weight}}. \tag{1}$$

As the sectoral inflation rates and the weights associated with them are random variables, we can therefore treat f_t as a functional random variable. It is often referred to as the cross-sectional density of sectoral inflation rates. In this regard, we can forecast the national (mean) inflation rate through (1).

Next, we define the fluctuation of the cross-sectional density, w_t, as a deviation from the well-defined unconditional expectation of the density, $\mathbb{E}[f_t]$:

$$w_t = f_t - \mathbb{E}[f_t], \ t = 1, ..., T. \tag{2}$$

We assume that this fluctuation disappears as the time period increases and this adjustment mechanism is specified by an autoregressive process. Hence, $\{w_t\}_{t=1}^T$ can be generated by an autoregressive process of order one in the functional space:

$$w_t = Aw_{t-1} + \epsilon_t, \ t = 1, \ldots, T, \tag{3}$$

where A is an autoregressive operator in the Hilbert space (H) and $\{\epsilon_t\}_{t=1}^T$ is the sequence of the functional white noise process. This model is referred to as the FAR of order one in a functional space, FAR(1) for short. Combining (2) and (3), we get:

$$f_t = \mathbb{E}[f_t] + Aw_{t-1} + \epsilon_t. \tag{4}$$

This portrays that the cross-sectional density at time t consists of the unconditional expectation $(\mathbb{E}[f_t])$

and the correction of the fluctuation at time $t-1$ (Aw_{t-1}).

First, at each time period, we estimate the cross-sectional density nonparametrically by the weighted kernel density estimator (Marzio and Taylor, 2004):

$$\hat{f}_t(z) = \frac{1}{h_t} \sum_{i=1}^{N} v_{it} K\left(\frac{z-\pi_{it}}{h_t}\right), \ t = 1, ..., T \tag{5}$$

where v_{it} is a time-varying sectoral weight, satisfying $\pi_t = \sum_i^N v_{it}\pi_{it}$ with $\sum_{i=1}^N v_{it} = 1$, K is a kernel, N is the number of sectors, and h_t is a bandwidth. One important issue lies in the selection of the appropriate kernel and the corresponding bandwidth. Here we opt to employ the Gaussian kernel, which is most popular in empirical studies. An optimal bandwidth is then derived by minimising a loss function and applying the cross validation selector. In this regard, we follow the Silverman's (1986) rule of thumb, which is the optimal bandwidth for the Gaussian kernel given by $h_t = 1.06\sigma_t N^{-1/5}$, where σ_t is the standard deviation of the sectoral inflation rate (π_{it}) at time t. Given the sequence of the estimated cross-sectional density functions, $\left\{\hat{f}_t\right\}_{t=1}^T$, the sequence of fluctuation is estimated by $\hat{w}_t = \hat{f}_t - \bar{f}$, where $\bar{f} = \frac{1}{T}\sum_{t=1}^T \hat{f}_t$ is the consistent estimate of $\mathbb{E}[f_t]$.

The autoregressive operator, A, in (3) can be estimated theoretically by $A = C_0^{-1}C_1$, where C_0 and C_1 are the autocovariance operator of order 0 and order 1, respectively, (see Bosq, 2000; Park and Qian, 2007, 2012). But, the autocovariance operators are defined in the infinite dimension. To avoid this ill-posed inverse problem in practice, we should project the autocovariance operators into a finite ℓ-dimensional subspace by applying a functional principal component analysis, and then estimate A consistently in the ℓ-dimensional subspace, denoted by $\hat{A}_\ell$, e.g. Park and Qian (2007, 2012).

Now, we can evaluate an m-step ahead forecasts of the cross-sectional density function, denoted $\hat{f}_{T+m|T}$, by

$$\hat{f}_{T+m|T} = \bar{f} + \hat{A}_\ell^m\left(\hat{f}_T - \bar{f}\right), \ m = 1, 2, \ldots, M, \tag{6}$$

where $\hat{f}_T$ is the estimate of the cross-sectional density at time T. It is then straightforward to evaluate the m-step ahead forecasts of the national inflation rate by integrating out the cross-sectional density for x as follows:

$$\hat{\pi}_{T+m|T} = \int_{\mathbb{D}} x\hat{f}_{T+m|T}(x)\, dx, \tag{7}$$

where $\mathbb{D} \subseteq \mathbb{R}$ is the domain of function $\hat{f}$ and the integral operator is numerically approximated by the middle Riemann sum given the grid set (see Step 1 in Appendix for details). It is clear from (7) that, when forecasting the national inflation rate, we do not need to forecast sectoral inflation rates and the corresponding weights, separately. We only need to forecast the functional variable, that is, the cross-sectional density.

Finally, we generate the empirical distribution of the forecasted national inflation rates $\left(\hat{\pi}_{T+m|T}\right)$ at each horizon, $m = 1, ..., M$, using the nonparametric bootstrap technique as follows: We first estimate (3), obtain residuals, $\hat{\epsilon}_t$, and collect the the pool of residuals, $\{\hat{\epsilon}_1, ..., \hat{\epsilon}_T\}$. Then, we draw the bth bootstrap sample of the prediction errors, $\left\{\hat{\epsilon}^{(b)}_{T+1|T}, \ldots, \hat{\epsilon}^{(b)}_{T+M|T}\right\}$ with replacement from the pool, and resample the m-step ahead forecast of the deviation by

$$\hat{w}^{(b)}_{T+m|T} = \hat{w}_{T+m|T} + \sum_{i=1}^{m} \hat{A}_{\ell}^{m-i} \hat{\epsilon}^{(b)}_{T+i|T}, \;\; m = 1, ..., M, \;\; b = 1, ...B, \tag{8}$$

where $\hat{w}_{T+m|T} = \hat{f}_{T+m|T} - \bar{f}$. The m-step ahead forecast of the cross-sectional density is then resampled by

$$\hat{f}^{(b)}_{T+m|T} = \hat{w}^{(b)}_{T+m|T} + \bar{f}, \quad m = 1, ..., M, \;\; b = 1, ...B, \tag{9}$$

which provides the (resampled) m-step ahead forecasts of the national inflation rates by

$$\hat{\pi}^{(b)}_{T+m|T} = \int_{\mathbb{D}} x \hat{f}^{(b)}_{T+m|T}(x)\, dx, \quad m = 1, ..., M, \;\; b = 1, ...B. \tag{10}$$

Then, the density forecast of the national inflation rate can be obtained from the empirical distribution of $\left\{\hat{\pi}^{(b)}_{T+m|T}\right\}_{b=1}^{B}$. In Appendix we provide the estimation algorithms for the semiparametric FAR model, the bootstrap-based inference as well as the robust modifications to address the issues related to the presence of structural breaks or the extreme events such as the global financial crisis.

3 Data and Analysis of Time Varying Moments

We use a set of sectoral inflation rates (defined as the annual percentage change) based on Consumer Price Index (CPI, 2005 =100) and their respective weights from the Office for National Statistics. In our case, sector implies the basket of goods and services that enter in the construction of CPI. The data spans over January 1997 to September 2013, thus giving a total of 201 monthly observations. In some sectors we have less observations as the disaggregation has started at a later date. Sub-sector consists of 79 sectors (January 1997 - November 2000), 84 sectors (December 2000 - November 2001) and 85 sectors (December 2001 - September 2013). The mean inflation rate in our sample (measured across sectors and over time periods) stands at 2.1% while the median is 1.8%. We observe that mean inflation rate is considerably higher at 3.1% in the post-crisis period (2008-2013) compared to 1.5% in the pre-crisis period (1997-2007). Both domestic (e.g., changes in value-added tax) and international factors (e.g., high oil price coupled with high food prices, rising demand from emerging economies, and fall in the value of sterling) are responsible behind this hike. Considerable heterogeneity exists across the sectors: Liquid Fuels experiencing the highest inflation (92.2%) and Information Processing Equipment the lowest (-40.2%). The standard deviation ranges from 0.62% (Restaurants & Cafes) to 27.8% (Liquid Fuels). Out of the 85 sub-sectors, inflation is positively skewed for 58 sub-sectors. The Jarque-Bera statistic is significant at the 5% significance level for 61 sub-sectors and at the 1% significance level for 55 sub-sectors, providing a strong evidence against the normality.

To understand the time-varying nature along with the degree of association between the first four cross-sectional moments, we provide their time series plots evaluated from the cross-sectional density of sectoral inflation rates (Figure 1). The maximum mean was in September 2008 (5.3%) and the minimum in August 2000 (0.3%). In 87 out of 201 months, inflation is positively skewed. Cross-sectional uncertainty of inflation substantially differs across time periods. It deviates 60 times by more than 1% from the 2% target over the full sample period (1997-2013), and 37 times occur in the post-crisis period (2008 - 2013). A simple paired t-test with unequal variance reveals that the cross-sectional mean inflation in pre-crisis period significantly differs from that of post-crisis period. We also note that cross-sectional RPV (RPS) on an average is 5.808 (-0.369) in the pre-crisis whereas 6.448 (0.362) in the crisis period.

[Figure 1 about here]

Figure 2 demonstrates the persistent nature of the first four cross-sectional moments for the whole sample, the pre-crisis and the post-crisis period. For the whole period and the pre-crisis samples, the first order moment is highly persistent. Interestingly, inflation uncertainty measured by the cross-sectional variance and skewness also exhibit high persistence for the whole and pre-crisis period. For the post-crisis period, however, RPV continues to be more persistent than the other three moments.

[Figure 2 about here]

We turn to analyse the correlation patterns among the cross-sectional moments. This may provide an insight to analyse the relationship between aggregate inflation and the standard deviation of relative price changes (often termed relative price variability, RPV) and between the mean and the skewness of relative price changes (henceforth relative price skewness, RPS). The seminal contribution by Friedman (1977) postulates that an increase in inflation uncertainty exerts a dampening effect on economic efficiency and possibly on output growth. Hence, we expect that there is a positive relationship among inflation rate, volatility and/or inflation uncertainty. Cukierman and Meltzer (1986) posit that higher inflation uncertainty increases the inflation rate while Holland (1995) predicts an opposite effect of uncertainty on inflation. Furthermore, the skewness could result from an exogenous form of downward nominal rigidity in product markets (Tobin, 1972) or endogeneity as suggested by the menu cost model (Ball and Mankiw, 1995). The former would imply a negative relation between the national inflation rate and RPS whereas the latter a positive one. This is important as the central banks need to differentiate between these two sources as they would exert different implications for an optimal monetary policy formulation. For example, with downward nominal rigidity, lower inflation rates are as harmful as they complicate the (downward) adjustment of relative prices whereas it is desirable in other case as it decreases the costs associated with changing prices.

Figure 3 displays the scatter plots between the pair of the cross-sectional moments. The pre-crisis period result in Panel 2 shows that the relationship between the mean and RPV and also between mean and RPS is U-shaped whist the RPV-RPS relationship appears approximately linear. For the post-crisis period (Panel 3) we observe that the INF-RPV relationship is still U-shaped, but the degree of association is weaker. On the other hand, the INF-RPS relationship becomes rather the inverse U-shaped whereas the sign of slope changes for the INF-RPK relationship from the pre-crisis to the post-crisis period. High inflation uncertainty in the post-crisis has a negative impact on the inflation rate perhaps due to the BOE's stabilizing policy towards the goal of long-run price stability (reduction of average of inflation) by applying tighter monetary policy to minimise the real costs of inflation uncertainty (Holland, 1995).

We investigate these relationships more formally using a simple two-regime threshold model, distinguishing between high and low inflation regimes according to whether current inflation exceeds the unconditional mean. Table 1 reports the results, showing that the impacts of RPV (RPS) on the mean inflation are measured at 0.145 and -0.106 (0.177 and -0.140), respectively, in the high- and the low-inflation-regime for the whole sample. For the pre-crisis period (column 2) the coefficients associated with RPV (RPS) are measured at 0.326 and -0.193 (0.288 and -0.184). But, the coefficients on RPV (RPS) are 0.085 and 0.044 (-0.022 and 0.286), in the high- and the low-inflation-regime, respectively, for the post-crisis period (column 3). Moreover, the null of symmetry is rejected for the INF-RPV and the INF-RPS relationship for the whole sample and the pre-crisis period, but not for the post-crisis period. On the other hand, the null of the symmetric mean-kurtosis relationship is strongly rejected for the whole and the post-crisis period, but not for the pre-crisis period. Interestingly, we find a positive association between standard deviation and skewness in the whole and pre-crisis period, though the association is stronger in the low-inflation regime.

[Figure 3 and Table 1 about here]

4 Forecasting the UK Inflation Rates

Our goal is to forecast the density of the national inflation rate utilising heterogeneous informational contents in sectoral inflation rates by combining the nonparametric kernel density estimation and the dynamic functional autoregressive modeling. Given our findings of persistent and interrelated moments of UK sectoral inflation rates, we need to accurately model these features to improve the performance of the density forecasting of the national inflation rate. This hybrid approach thus enables us to take into account the relative advantage of parametric and nonparametric approach in a robust manner as the fully parametric specification is inappropriate to unravel an exact relationship among higher-order cross-sectional moments.

Following Park and Qian (2007, 2012), we consider three basic models, namely FAR, AVE and LAST. For forecasting purpose AVE utilises the average of all observed distributions while LAST uses the last observation:

$$AVE: f_t = \bar{f} + \epsilon_t; \quad LAST: f_t = f_{t-1} + \epsilon_t. \tag{11}$$

In practice, however, forecasting performance may be significantly affected by the presence of structural breaks or by the fact that inflation and its moments are highly persistent. To address these important issues, we apply the intercept shift or the first difference modifications recommended by Clements and Hendry (2002, 2006), and consider the additional nine modified models, referred to as FAR3M, FAR6M, FAR9M, FAR12M, AVE3M, AVE6M, AVE9M, AVE12M and DFAR. See Appendix for further details.

4.1 Pseudo Out-of-sample Forecasting Evaluation

We conduct four exercises in a recursive manner to evaluate the forecasting performance of all the functional models. This practice of holding out sample is called "pseudo real time" (Elliot and Timmermann, 2008) or "quasi-real time recursive out-of-sample" experiments (Faust and Wright, 2012).

4.1.1 Forecasting of Cross-sectional Density Function

We conduct the evaluation of the cross-sectional density forecasting by comparing the divergence criteria such as the Hilbert norm (D_H), the uniform norm (D_U) and the generalised entropy (D_E) defined by

$$D_U\left(\hat{f}_t, f_t\right) = \frac{\int \left(\hat{f}_t\left(x\right) - f_t\left(x\right)\right)^2 dx}{\int \left(\hat{f}_t\left(x\right)^2 + \hat{f}_t\left(x\right)^2\right) dx},\ D_H\left(\hat{f}_t, f_t\right) = \frac{\sup_x \left|\hat{f}_t\left(x\right) - f_t\left(x\right)\right|}{\sup_x f_t\left(x\right)},\ D_E\left(\hat{f}_t, f_t\right) = \int \hat{f}_t\left(x\right) g\left(\frac{\hat{f}_t\left(x\right)}{f_t\left(x\right)}\right) dx, \tag{12}$$

where $\hat{f}_t$ (f_t) is the forecasted (realised) density function and $g\left(y\right) = \left(\gamma - 1\right)^{-1}\left(y^{\gamma} - 1\right)$ with $\gamma > 0$ and $\gamma \neq 1$. We set $\gamma = 1/2$ (e.g. Park and Qian, 2007). If g is a natural log function, D_E becomes the Kullback-Liebler divergence measure. All three quantities are non-negative and become zero if $\hat{f}_t = f_t$. D_H is useful for evaluating the goodness-of-fit of the model, D_U is informative for comparing the closeness of the function shape, and D_E assesses the difference in information contents between the forecasted and the true density function.

We first estimate a total of 12 FAR-based models introduced above over the period January 1997 - December 2002, and compute one month-ahead forecast of cross-sectional density function for January 2003. We repeat the process moving forward one month at a time in a recursive manner, ending with forecasts for September 2013 based on the estimated models over January 1997 - August 2013. This generates a total of 129 observations for evaluating the closeness between one-step-ahead cross-sectional density forecasts and the actual counterparts for each of the twelve models over the forecasting horizon January 2003 - September 2013. Further, to evaluate the closeness measure for the pre-crisis period (2003 - 2007), and the post-crisis period (2008 - 2013), we consider the subtotal of 60 and 69 observations, respectively.

Table 2 presents these evaluation results in terms of both mean and median. Overall we find that DFAR and LAST have the minimum distances for almost all cases. On the other hand, the AVE has the maximum distance for all criteria. But, modified FAR and AVE models have intermediate values with modified FAR models displaying significantly lower values.

[Table 2 about here]

4.1.2 Point Forecasting of National Inflation Rate

This section focuses on the forecasting of the national inflation rate. As benchmark models, we consider an AR model with a maximum lag, 12, and an ARCH-in-mean model (ARCH-M). These two models utilise only aggregate inflation rates. In contrast our proposed models utilise the disaggregate sectoral inflation rates. We estimate models from January 1997 to December 2002, compute twelve-month-ahead forecasts of the national inflation rate by (10), and recursively repeat the process moving forward one month at a time. Thus, this exercise ends with forecasts for October 2012 - September 2013, based on models estimated over January 1997 - September 2012. We obtain 118 observations from each of m-month-ahead forecast, $m = 1, \ldots, 12$, giving a total of 1,416 experiments. Similarly, we obtain the subtotal of 588 and 696 experiments respectively for the pre- and the post-crisis periods.

We evaluate the forecasting performance of the FAR models by comparing forecasting errors quantitatively and statistically. Quantitatively, we compare an absolute forecasting error loss by mean absolute error (MAE):

$$MAE = N^{-1} \sum_{t=1}^{N} |\hat{\pi}_t - \pi_t| \,. \tag{13}$$

Further, we statistically test the equality of forecasting accuracy using the testing procedure developed by Diebold and Mariano (1995) (hereafter, DM test) based on MAE. (Notice that we have obtained qualitatively similar results when using the mean square error, which are not reported here to save space.)

First, we compare MAE for a total of fourteen models: twelve functional models and two benchmark models respectively for the whole sample, the pre- and the post-crisis period. Table 3 reports the results for the whole sample. We find that DFAR and LAST have the smallest values compared to the AR and ARCH-M model for all the forecasting horizons (except for the 1-month ahead forecast compared to the ARCH-M model). The performance of AVE is worse than the AR model for all forecast horizons and that of the ARCH-M model for the first six months. The intercept shift models of FAR and AVE (FAR3M-FAR12M and AVE3M-AVE12M) have intermediate values: the MAE is lower compared to the benchmark AR model for all the forecast horizons except for 2-3-month horizons for AVE9M and AVE12M. For the pre-crisis period, the superior performance of DFAR is again pronounced, see Table 4. The modified FAR and the AVE models dominate the AR and ARCH-M models for almost all of the forecasting horizons. Furthermore, from Table 5, we observe qualitatively similar results for the post-crisis period. The MAEs for FAR3M, AVE3M, DFAR and LAST models are significantly lower than those for AR and ARCH-M models.

Next, we formally test the equality of forecasting accuracy of each of functional models against the benchmark models using the DM test. Specifically, we test the validity of the null hypothesis that the MAE of the functional model is equal to that of benchmark model against the alternative that the former is less than the latter. The DM statistic follows the standard normal distribution under the null asymptotically. To evaluate it, we have to estimate the long-run variance of the difference between two forecasting error losses. Here we employ the Newey and West (1987) method with Bartlett window, denoted $w_{j,T} = 1 - j/(q_T + 1)$, where q_T equals to the integer part of $4(T/100)^{2/9}$. For the whole sample, we confirm that DFAR and LAST (and other modified functional models such as FAR3M and AVE3M) convincingly beat the AR at the short forecasting horizons and the ARCH-M for all horizons except for one month ahead. In particular, the pre-crisis sample portrays the superior performance of the DFAR model against the benchmark models for all forecast horizons, though the post-crisis gains are much less clear than those for the pre-crisis.

[Tables 3, 4 and 5 about here]

The results reported in Tables 3, 4 and in 5 clearly demonstrate the superior performances of the modified FAR models, especially DFAR, in forecasting the national inflation rates. The models incorporating the intercept shift modifications, such as the FAR3M and the AVE3M, also outperform the AR and ARCH-M models. By contrast, the performance of (unmodified) FAR and AVE models is rather poor and dominated even by the AR model. This clearly highlights an importance of applying the proposed modifications to improve the forecasting performance of the functional model in practice. Combining the forecasting evaluation results for both the cross-sectional density and the mean national inflation together, we come to a conclusion that the use of aggregate data is likely to result in misleading forecasts as it ignores the underlying dynamics of the heterogeneous micro units, (Pesaran and Smith, 1995; Hsiao et al., 2005).

4.1.3 Density Forecasting of National Inflation Rate

In this section we evaluate the density forecasting performance of national inflation rate for each of the twelve functional models, using the probability integral transformations (PIT) proposed by Diebold et al. (1998) and widely employed in the literature (e.g., Clements, 2004; Mitchell and Hall, 2005). Table 6 reports the Kolmogorov-Smirnov (KS) test results for testing the null hypothesis that the forecasted and the actual density functions are equal. Using the same recursive estimation procedure as in Subsection 4.1.2, we have obtained the total of 129 PIT observations for the full forecasting period (2003 -2013) in Column 1. To construct the PITs associated with the density forecasts of national inflation we employ the bootstrap schemes as described in Step 4 in Appendix. Similarly, we report the test results in Columns 2-4 for the three subperiods using 60 PIT observations over the pre-crisis period (2003-2007), 48 PIT observations over the crisis period (2008-2011), and 45 PIT observations over the post-crisis period (2010-2013). Moreover, in order to mitigate the potentially detrimental impact on the KS test results of the crisis period and investigate how the functional models perform during the post-crisis period, we now analyse the crisis (2008-2011) and the post-crisis (2010-2013) period, separately. In doing so the overlapping of 2010 and 2011 is made inevitable due to the minimum sample requirement for evaluating the KS test statistics.

We conduct the analysis using both the standard bootstrap with $h = 1$ and the robust bootstrap with $h = 3$ (see (A10) in Appendix for details). Panel A in Table 6 reports the test results obtained with $h = 1$. We find that the null hypothesis is not rejected only for LAST and AVE3M over the full forecasting period. Turning to the three subperiods, we establish that the null is not rejected for modified FAR models (e.g. FAR3M, AVE3M and DFAR) as well as for LAST during both the pre- and the post-crisis period. During the crisis period, however, the null hypothesis is strongly rejected for all the models. This may not be surprising as the empirical density evaluated by the standard bootstrap is likely to severely underestimate the realised ones, especially in the presence of extreme events such as the global financial crisis. To explicitly address this issue we next examine the test results with robust bootstrap presented in Panel B. The overall performance of the PIT test is shown to be improved with the robust bootstrap. The null is not rejected for most modified FAR models over the full forecasting period as well as during the pre- and the post-crisis period. During the crisis period, however, the null is not rejected for DFAR only. This also displays an importance of applying the robust bootstrap modification.

[Table 6 about here]

In sum our results as reported in Tables 2, 3, 4, 5 and in Table 6 reveal that DFAR, FAR3M, AVE3M and LAST generally outperform the other models in forecasting both the cross-sectional density of sectoral inflation rates and the density of the national inflation rate.

Next, we evaluate the accuracy of the interval forecasting for these four better performing functional models (FAR3M, AVE3M, LAST and DFAR) by constructing an interval from the empirical distribution given the coverage probability, 95%. To this end we apply the conditional coverage test (hereafter, CC test) developed by Christoffersen (1998). The CC test uses an indicator function taking unity for the case that a realised national inflation rate is covered by the interval for each of forecasting horizons, and 0 otherwise. Then, it tests if the conditional expectation of the binary random variable generated by the indicator function is equal to the coverage probability. Christoffersen (1998) shows that it is equivalent to testing if the sequence of the binary random variable is identically and independently distributed Bernoulli with parameter p, coverage probability. Hence, the LR statistic simultaneously tests if the unconditional coverage probability is p (unconditional coverage test) and the the binary random variable is independent (independent test). It follows the chi-squared distribution with two degrees-of-freedom under the null hypothesis.

As with the PIT-based tests, we construct the empirical distribution of the national inflation rate using the same bootstrap schemes as described above, and report the test results with the standard bootstrap ($h = 1$) and with the robust bootstrap ($h = 3$) respectively in Panel A and B of Table 7. In turn, each panel contains the three sub-panels corresponding to the pre-crisis period (January 2003-December 2007), the crisis period (January 2008-December 2011) and the post-crisis period (January 2010-September 2013).

For the pre-crisis period we estimate models from January 1997 to December 2002, compute twelve-month-ahead forecasts of cross-sectional density, and recursively repeat the process moving forward one month. This exercise ends with forecasts for the period, January 2007 - December 2007, based on models estimated over January 1997 - December 2006. This produces 48 observations from each of m-month-ahead forecast, $m = 1, \ldots, 12$, giving a total of 576 experiments. The pre-crisis period results reported in Panel A.1 of Table 7 show that the CC test is not rejected for DFAR and LAST for all forecasting horizons whereas it is not rejected for AVE3M and FAR3M up to 6 - 8 month ahead forecasts. Turning to Panel B.1 of Table 7 that reports the pre-crisis period test results with the robust bootstrap, we find that the CC tests do not reject the null hypothesis for FAR3M, DFAR, and LAST for all forecasting horizons.

For the post-crisis period, we conduct two experiments. First, we start estimating models from January 1997 to December 2007, and compute twelve-month-ahead forecast of cross-sectional density, recursively. We thus obtain 36 observations from each of m-month-ahead forecast, $m = 1, \ldots, 12$, giving a total of 432 experiments. As expected, all models are convincingly rejected for all forecasting horizons even with the robust bootstrap during this crisis period (2008 - 2011) (see Panels A.2 and B.2 in Table 7). Second, we start estimating models from January 1997 to December 2009, compute twelve-month-ahead forecast of cross-sectional density in the same recursive manner, and obtain 33 observations from each of m-month-ahead forecast, $m = 1, \ldots, 12$, giving a total of 396 experiments. The test results with the standard bootstrap reported in Panel A.3, indicate that the CC test is not rejected only for the shorter horizons. The performance of the CC test is improved with the robust bootstrap, especially for DFAR and LAST models, both of which are not rejected for almost all forecasting horizons (see Panel B.3 of Table 7).

[Table 7 about here]

Given the results for the cross-sectional distribution, mean inflation, density evaluation and interval forecasting performance, we are able to justify the statistical validity of our modified FAR modeling approach using the disaggregate data even in the presence of the crisis period (albeit not perfectly).

4.2 Fan Chart Analysis and Probability Event Forecasting

Given the success with forecasting evaluation exercises, we extend to augment the forecasting exercise in two important ways: the national inflation rate forecasting with uncertainty bands, similar to the fan-chart provided by the BOE quarterly inflation report available from the BOE's website (publications/inflationreport/irfanch.htm), and the probability event forecasting of inflation targets. Some central banks adopting the inflation-targeting tend to describe the uncertainty relating to their forecasts verbally. However, since its introduction by BOE in 1996, the majority of the central banks convey the forecast uncertainty by fan charts. Fan-charts represent the most likely future development and thus help to improve communication amongst practitioners and policymakers by laying more emphasis on the risks of the inflation forecast and their direction.

4.2.1 Fan Chart

We provide fan-charts for five different periods: January 2003-December 2004, January 2005-December 2006, January 2007-December 2008, January 2009-December 2010, and January 2011-December 2012 for the four better performing models (FAR3M, AVE3M, DFAR and LAST). To construct these figures, for example, for the first period (January 2003-December 2004), we estimate the four functional models over January 1997-December 2002, and compute one- to twenty-four-month-ahead forecasts of mean inflation rates. The confidence bands are then evaluated using the bootstrap schemes employed in Section 4.1. We apply the same method to other sample periods for obtaining the 24-month-ahead forecasts of mean inflation rates with confidence bands, though we have used all the available observations in constructing the conditional forecasts, e.g., we estimate the models over January 1997-December 2010 for the last forecasting period (January 2011-December 2012). Here, we only report the results for the robust bootstrap ($h = 3$) in Figure 4 to save space, though the results for the standard bootstrap ($h = 1$) are qualitatively similar. In each one of these figures, the dash-line presents in-sample national inflation forecast and the solid line shows the actual national inflation. The confidence bands (shades) present 12 intervals from 2.5th percentiles (bottom) to 97.5th percentiles (top) out of empirical distribution $\left\{\hat{\pi}^{(b)}_{T+m|T}\right\}_{b=1}^{B}$ for each horizon, $m = 1, ..., 24$.

We find from the first and the second columns of Figure 4 that realised inflation is well within the confidence band with forecasted mean inflation tracking actual one reasonably well. The mean is being projected at 1.64%, 1.30%, 1.35% and 1.57% respectively during January 2003-December 2004 for FAR3M, AVE3M, LAST and DFAR models. The band widens with the horizon, indicating increasing uncertainty about inflation. The fan-chart for the January 2005-December 2006 period (especially for DFAR) is strikingly similar with the one published by the BOE quarterly bulletin in fourth quarter of 2004 for the next 24-months.

The third column presents the results for the crisis period (January 2007-December 2008), and depicts that actual inflation rates lie outside the band only during July-October in 2008. This period reflects the intensification of the subprime mortgage crisis followed by the Lehman Brother's bankruptcy in September 2008. During this period inflation expectations had drifted upwards as market participants viewed monetary authorities focusing mainly on softening the impact of financial instability on the real economy and less on the commitment to fighting inflation. The projected mean inflation of 24-month-ahead forecast is stable at 2.59%, 2.74%, 2.60% and 2.62% for FAR3M, AVE3M, DFAR and LAST, with the uncertainty band lying between 0.5% to 4.35% for the DFAR model. This shows greater uncertainty surrounding the point forecasts. Next, during the forecasting period (January 2009-December 2010, the fourth column), we observe that actual inflation rate was surprisingly low at 0.98% in September 2009, mainly due to deflation scare with the forthcoming launch of the unprecedented policy of asset purchases financed by central bank, the so-called Quantitative Easing (QE). This observation still lies within the band of DFAR and LAST models, showing that the generated uncertainty band from our proposed models is also able to incorporate this extreme low inflation. The BOE quarterly bulletin in the third quarter of 2011 reports that the QE may have increased inflation by 0.75% to 1.5%, and these were economically significant. We believe that our fan-chart analysis also tracks this increase quite well. The final column displays the results for the last forecasting period (2011-2012), and shows that the actual inflation rates lie well within the confidence band for all the four models. Notice that the in-sample national inflation rate forecast (dash-line) is always above the target of 2% inflation which is line with the opening remarks at Press Conference in November 2010 by the Governor of BOE stating that inflation is likely to remain above the 2% target throughout 2011. We also note the fan-charts in this period especially from the FAR3M and DFAR models is remarkably similar with that of BOE fan-charts published in November 2010 report.

[Figure 4 about here]

In sum, the fan-chart analysis seems to support our modeling approach in a qualitative sense and thus provide in probability terms the uncertainty inherent to any point forecast. The forecasting exercise highly depends on the state of the economy, especially during the crisis. Our framework thus acts as a complementary tool for generating the density forecast for inflation in a flexible manner without the inherent subjectivity associated with the fan-charts reported by the BOE.

4.2.2 Probability Event Forecasting

We examine the probability of achieving the target of $\pi < 2\%$ (thereafter referred to as the first target) and $1\% < \pi < 3\%$ (the second target) as announced by BOE in its inflation targeting framework using the probability forecasting results for the four models (FAR3M, AVE3M, DFAR and LAST) along with the actual national inflation for the same five different periods as analysed above. This is an interesting exercise because many economic agents are likely to judge the performance of the central bank in relation to realised outcomes rather than largely unobservable forecast outcomes. Hence, one may think of our inflation forecasts as predictions of the public perception of the success or failure of the central bank in achieving inflation stabilization. Here we have also used the same recursive estimation and bootstrap scheme.

Figure 5 shows that the probability of achieving the first target is large (around 70%) for the first two pre-crisis forecasting periods (2003-2004 and 2005-2006). With the second target, the probabilities decrease slowly with the horizon but stay around 65-70%. For the crisis period (2007-2008), the probability of keeping the first target increases slowly as the horizon rises, but stays fairly low (25% for DFAR) over the two-year horizon; whereas the probability of keeping the second target almost remains at 60% (see the third column). The actual data reveals that during this period, the two targets of $\pi < 2\%$ and $1\% < \pi < 3\%$ are violated 88% and 42% of times, respectively. Therefore, we believe that our probability event forecasting exercise is able to capture the violations almost accurately. The fourth period (2009-2010) analysis shows that the probability of keeping the first target is fairly low (10%-16%) whereas for the second target the average probability remains around 40% for DFAR. Over this period, the two targets are violated 83% and 54% of times, respectively. For the last period (2011-2012), the average probability of keeping the first target ranges from around 5% (FAR3M) to 10% (DFAR). Turning to the second target, we observe that the probability increases slowly with horizon but remains low from 20% (AVE3M) to 35% (DFAR). The actual data reveals that the first target is violated always whereas the second target is being violated 67% of times, respectively.

[Figure 5 about here]

In the analysis of the probability event forecasts, it would be good to complement the discussion with more formal measures of probability forecast performance, such as probability score measures (e.g. quadratic probability score or QPS; Brier, 1950) and assessment of the probability forecasts' calibration, e.g. Diebold and Lopez (1996) and Lopez (2001). QPS tracks the average squared divergence between the probability forecast and the realised outcome of an event with the smaller value indicating more accurate forecast. There are alternative scoring rules such as logarithmic and spherical probability scores. The calibration of probability forecasts is defined as the degree of equivalence between the forecasted probabilities and observed frequencies within subsets of the unit interval. Under the null hypothesis of their equivalence, the calibration test statistic asymptotically converges a standard normal distribution (Seillier-Moiseiwitsch and Dawid, 1993).

Table 8 and 9 report the probability scores (QPS) and the associated calibration test results, respectively. We conduct this analysis for three periods: the full-sample (2003-2013), the pre-crisis (2003-2007) and the post-crisis (2008-2013). For the full-sample we estimate models from January 1997 to December 2002, compute 24-month-ahead forecasts of cross-sectional density, recursively. This exercise ends with forecasts for the period, October 2011-September 2013, based on models estimated over January 1997-September 2011. This produces 106 observations for evaluating QPS and the calibration test statistics for each of m-month-ahead forecasts, $m = 1, \ldots, 24$. Similarly, we obtain 37 and 46 test sample observations for the pre-crisis and the post-crisis.

[Tables 8 and 9 about here]

Consider the results for the first event ($\pi < 2\%$). Over the full-sample period (2003-2013), DFAR produces the lowest QPS over the shorter horizons (up to 6 month), followed by AVE3M and FAR3M over the longer horizons. Turning to the calibration test results in Table 9, we find that the null hypothesis is not rejected for all the four models over the shorter horizons (up to 2 to 4 months), though the null is rejected over the longer horizons. This mainly reflects uncertainty significantly growing with the forecast horizon as is evident with the fan-chart analysis in Figure 4. The pre-crisis period (2003-2007) analysis shows a rather mixed result: QPS is the smallest for DFAR over 1 month horizon, for AVE3M over 2-3 month horizons and for FAR3M over 4-12 month horizons. Despite growing uncertainty, the null of the calibration test is not rejected for all four models over the entire forecasting horizons except for AVE3M over the longer horizons (7-12 months). For the post-crisis period (2008-2013), DFAR produces lower probability scores over the shorter horizons (up to 6 month horizons), and then AVE3M over the remaining horizons. The calibration test results confirm that the null is not rejected for all four models only over the shorter 1-2 month horizons. However, the null is not rejected for AVE3M over the entire horizons.

Next, we turn to the analysis of the second event ($1\% < \pi < 3\%$). Over the full-sample (2003-2013), DFAR produces lower probability scores over the first 9-month horizons and then FAR3M over 10-12 months. The calibration test results reveal that the null is not rejected for all four models over the short horizons (up to 6 month) except for AVE3M over 1 month horizon. In particular we do not reject the null for DFAR over the entire horizons. For the pre-crisis period we obtain the similar patterns of QPS to the full sample case: QPS is the lowest for DFAR over the first 7-month horizons, and then for FAR3M over the remaining horizons. However, the calibration tests reject the null for all four models over the whole forecasting horizons except for FAR3M and DFAR marginally over 1-month horizon. This strong rejection may reflect the extreme difficulty associated with forecasting rare events (the actual frequency is 97% during the pre-crisis period), especially when forecasting uncertainty grows significantly with the horizon. Nevertheless, it is still encouraging to find out that the four models are able to produce the probability forecasts higher than 90% over the shorter horizons up to 3-month. Finally, for the post-crisis period (2008-2013), DFAR produces the smaller QPS over the entire horizons, except over 2 month horizon for LAST and over 10-11 month horizons for FAR3M. Moreover, the null of the calibration test is not rejected for all four models over the shorter (up to 3 month) and the longer (8-12 month) horizons.

Combined together, we find that both QPS and calibration test results provide an overall support for modified functional models, especially the satisfactory performance of DFAR over the short horizons. In particular, we note that DFAR produces the smallest QPS in 38 times and then FAR3M in 21 times out of a total of 72 cases combining the two event forecasts.

Our results from the fan-chart and probability event forecasting exercise reveals: we are able to provide a complementary tool for generating the complete density forecast for inflation in a flexible manner and analyse the performance of the central bank in achieving announced inflation target. We conclude that a thorough examination of the time-varying cross-sectional distribution provides satisfactory performance based on pseudo out-of-sample forecasting exercises. Clements (2004) provides evidence that the MPC current and next quarter forecasts performs better based on mean squared forecast error compared to no-change forecasts (based on Gaussian density with mean given by the actual inflation rate in the last period), but not the year-ahead density forecasts. Our proposed model offers more in this direction. We can forecast the mean inflation rate as well as the probability event forecasting of inflation targets. The results from the pseudo out-of-sample forecasting exercise of cross-sectional density, mean inflation, density evaluation, and interval forecasting indicate the statistical validity of our modified FAR modeling approach using the disaggregate data even in the presence of the crisis period. On the other hand, the fan-chart and probability event forecasting exercise provides a complementary tool for generating the complete density forecast for inflation without the inherent subjectivity associated with the BOE's fan-chart analysis.

5 Conclusion

In this paper, we use the semi-parametric functional autoregressive approach to model the time-varying distribution of the UK monthly inflation rate over the period January 1997-September 2013 using sectoral CPI data. Our approach exhibits several novel features. First, our model is parsimonious in nature and provides additional information in forecasting the national inflation rate and the associated uncertainty by allowing for (rather complex) cross-sectional dependence and exploiting the variation across sectors. Second, our approach can easily provide the time-varying cross-sectional moments obtained at each period which enables us to analyse the stylised descriptive features amongst all the moments. Third and importantly, our nonparametric modeling of the time-varying density of inflation rates across sectors does not impose any assumptions on the structure of the distribution or on the number of cross-sectional units. Finally, our approach is much simpler as compared to the large-scale simultaneous equation macroeconometric models developed by the Fed and the BOE and the system-based cointegrating VAR approach of Garratt et al. (2006), and thus provides a complementary tool for generating the complete density forecast for inflation in a flexible manner. Thus, our framework provides an alternative for use by the independent researchers to the forecasts made by the central bank, and allows us to incorporate complex dynamic responses by the policymakers to the disaggregate sectoral shocks.

There is broad agreement that high inflation is qualitatively different from low and stable inflation while at the same time there are concerns about excessively low inflation due to the implications of the zero lower bound on nominal interest rates for the conduct of monetary policy (e.g., Greenwood-Nimmo et al., 2012). The use of the semi-parametric FAR approach in forecasting the density of inflation rates clearly helps us to evaluate the probability forecasting of inflation lying within a particular range at any given horizon. This will naturally provide valuable information for policy makers in advance of their rate setting decisions.

Our findings have an important policy implication as an inflation targeting central bank may face a trade-off between low inflation on the one hand and increased volatility and (absolute) skewness on the other hand. The policy reaction of the central bank would therefore differ depending on the nature of the shock and the regime. A shock originating in the high RPV sector may shift the distribution of this sector rightward (leftward) and thereby would increase (decrease) the mean inflation. Here, the low RPV sector could follow the distribution of the high RPV sector due to wage indexation. With a rightward shift of the distribution, to maintain overall low inflation, the central bank would follow an aggressive policy intervention to prevent the movement of the low inflation sector towards the high inflation sector. However, if the distribution of this variance sector shifts leftward, the central bank may remain passive in terms of policy intervention as long as it does not create a deflationary environment. As inflation targeting monetary policies are usually set with recourse to inflation forecasts, our proposed work on forecasting the time-varying distribution of sectoral inflation rates may provide policymakers with an invaluably enriched information set.

Although our work is limited to the UK data, this approach can be easily applied to any country with comparable data. A few issues remain to be resolved for future research. Firstly, to deal with the structural breaks and/or regime switching events, one can construct a mixed model; first, the slowly-varying local trend level of inflation rates are modeled separately using the regime-switching type models or obtained as the long-run survey expectations interpreted as representing agents' perceptions of the long-run inflation target; then we apply the FAR modeling to the filtered residuals, and forecast uncertainty by bootstraps. Secondly, our proposed methodology can easily be applied to the SPF data or any other survey data such as published by Livingston Survey or Consensus Economics Incorporation or even to the BLS data. Finally, one can extend our framework in a multi-dimensional context by adding some other macroeconomic and financial variables. One feasible approach is to combine the marginal distributions of each series forecasted by the FAR model using the Copula approach. This would enable one to perform several multi-dimensional tasks, which would enhance our understanding of a range of macroeconomic stylised facts, and provide us an invaluably enriched information set to improve forecasting performance.

Appendix: Estimation and Forecasting Procedure in Details

We describe a detailed estimation and forecasting procedure for national inflation rates using the FAR approach. Disaggregate inflation rates, π_{it}, consist of N sectors over T periods. For simplicity, we consider the following balanced panel data (as highlighted in Introduction, our approach can be easily applied to the unbalanced panels):

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{21} & \cdots & \pi_{N1} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{1T} & \pi_{2T} & \cdots & \pi_{NT} \end{bmatrix}_{T\times N}.$$

Step 1: The kernel density estimation of cross-sectional density of inflation rates We set a $p \times 1$ vector of grids, $\mathbf{z} = (z_1, \ldots, z_p)'$, which covers the range of realised inflation rates. In an empirical analysis we set $p = 1,024$. We estimate a cross-sectional density of inflation by the weighted kernel density estimator:

$$\hat{g}_t(z_j) = \frac{1}{h_t}\sum_{i=1}^{N} v_{it} K\left(\frac{z_j - \pi_{it}}{h_t}\right),\ j = 1, \ldots, p;\ t = 1, ..., T, \tag{A.1}$$

where K is a kernel, N the number of sectors, and h_t a bandwidth. Following Silverman (1986), we use a Gaussian kernel with an optimal bandwidth given by $h_t = 1.06\hat{\sigma}_t N^{-1/5}$, where $\hat{\sigma}_t$ is the cross-sectional standard deviation of π_{it}. Then, we normalise the estimated density function by

$$\hat{f}_t(z_j) = \frac{\hat{g}_t(z_j)}{RSUM(\hat{\mathbf{g}}_t)}, \tag{A.2}$$

where $RSUM(\hat{\mathbf{g}}_t)$ is a numerical middle Riemann sum of $\hat{\mathbf{g}}_t = (\hat{g}_t(z_1), \ldots, \hat{g}_t(z_p))'$ given by

$$\begin{aligned} RSUM(\hat{\mathbf{g}}_t) &= \frac{1}{2}\left[\sum_{j=1}^{p-1} \hat{g}_t(z_j)(z_{j+1} - z_j) + \sum_{j=2}^{p} \hat{g}_t(z_j)(z_j - z_{j-1})\right] \\ &= \frac{1}{2}\left[\sum_{j=1}^{p-1} \hat{g}_t(z_j) + \sum_{j=2}^{p} \hat{g}_t(z_j)\right]\Delta,\ \Delta = z_j - z_{j-1} \text{ for } j = 2, \ldots, p. \end{aligned}$$

Next, we construct the matrix of the cross-sectional densities,

$$F = \left[\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \ldots, \hat{\mathbf{f}}_T\right] = \begin{bmatrix} \hat{f}_1(z_1) & \hat{f}_2(z_1) & \cdots & \hat{f}_T(z_1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{f}_1(z_p) & \hat{f}_2(z_p) & \cdots & \hat{f}_T(z_p) \end{bmatrix}_{p\times T}.$$

We then estimate an unconditional mean function by

$$\bar{\mathbf{f}} = \begin{bmatrix} T^{-1}\sum_{i=1}^{T}\hat{f}_i(z_1) \\ \vdots \\ T^{-1}\sum_{i=1}^{T}\hat{f}_i(z_p) \end{bmatrix}_{p\times 1}, \tag{A.3}$$

and generate the matrix of fluctuations around the unconditional mean function by

$$W = [\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_T] = \begin{bmatrix} \hat{w}_1(z_1) & \hat{w}_2(z_1) & \cdots & \hat{w}_T(z_1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_1(z_p) & \hat{w}_2(z_p) & \cdots & \hat{w}_T(z_p) \end{bmatrix}_{p\times T}, \ \hat{\mathbf{w}}_t = \hat{\mathbf{f}}_t - \bar{\mathbf{f}}.$$

Step 2: The estimation of FAR model We estimate an autocovariance operators of order 0 and 1 by

$$\hat{C}_0 = T^{-1}\sum_{t=1}^{T}\hat{\mathbf{w}}_t\hat{\mathbf{w}}_t', \ \hat{C}_1 = (T-1)^{-1}\sum_{t=2}^{T}\hat{\mathbf{w}}_t\hat{\mathbf{w}}_{t-1}'.$$

Then, we estimate the pair of eigenvalue (λ_j) and eigenfunction $(\mathbf{v}_j)$ of $\hat{C}_0$ by

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix}_{p\times 1}, V = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_p \end{bmatrix} = \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{p1} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1p} & v_{2p} & \cdots & v_{pp} \end{bmatrix}_{p\times p}.$$

We choose ℓ eigenvalues, $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$, and eigenfunctions, $(\mathbf{v}_1, \ldots, \mathbf{v}_\ell)$, and then approximate the inverse of $\hat{C}_0$ in an ℓ-dimensional subspace by

$$\hat{C}_{0,\ell}^{+} = \sum_{k=1}^{\ell}\lambda_k^{-1}\mathbf{v}_k\mathbf{v}_k'. \tag{A.4}$$

The choice of ℓ is guided by applying a functional principal component analysis (FPCA) and a cross validation (CV). FPCA explains the variation of the fluctuation and CV chooses an optimal dimension

$\ell\,(\leq \ell_{max})$ by minimising the following criterion:

$$\sum_{i=1}^{N_{cv}} RSUM\left(\left[\hat{\mathbf{w}}^{\ell}_{T-i+1} - \hat{\mathbf{w}}_{T-i+1}\right]^2\right),$$

where N_{cv} is the number of the last observations used in CV and $\hat{\mathbf{w}}^{\ell}_{T-i+1}$ the in-sample forecasts of $\mathbf{w}_{T-i+1}$ on the ℓ-dimensional subspace. In our empirical analysis we set $\ell_{max} = 20$, and find that CV selects the optimal value of ℓ ranging between 5 and 10. The autoregressive operator A in (3) is now consistently estimated in the ℓ-dimensional subspace by

$$\hat{A}_{\ell} = \hat{C}_1 \hat{C}^{+}_{0,\ell}.$$

Step 3: The density forecasting of national inflation rates An m-step ahead forecasts of the cross-sectional density is evaluated by

$$\hat{\mathbf{f}}_{T+m|T} = \bar{\mathbf{f}} + \hat{A}^{m}_{\ell}\hat{\mathbf{w}}_T \text{ for } m = 1, ..., M,$$

where $\hat{\mathbf{w}}_T$ is the estimate of the fluctuation at time T. Next, the m-step ahead forecasts of the national inflation are obtained by integrating out the cross-sectional density for $\mathbf{z}$ as follows:

$$\hat{\pi}_{T+m|T} = RSUM\left(\mathbf{z} \odot \hat{\mathbf{f}}_{T+m|T}\right) \text{ for } m = 1, \ldots, M,$$

where $\odot$ stands for an element-by-element multiplication operator.

To construct the empirical density forecasts of national inflation rates, $\hat{\pi}_{T+m|T}$ for $m = 1, ..., M$, we conduct the nonparametric bootstrap by drawing the prediction errors, $\left\{\hat{\epsilon}^{(b)}_{T+1|T}, \ldots, \hat{\epsilon}^{(b)}_{T+M|T}\right\}$ for $b = 1, \ldots, B$, with replacement from the pool of the sample residuals, $\{\hat{\epsilon}_1, \ldots, \hat{\epsilon}_T\}$, where $\hat{\epsilon}_t = \hat{\mathbf{w}}_t - \hat{A}_{\ell}\hat{\mathbf{w}}_{t-1}$. Then, the m-step ahead forecasts of the fluctuation and of the cross-sectional density are re-sampled respectively by (8) and (9), which we denote by $\hat{\mathbf{w}}^{(b)}_{T+m|T}$ and $\hat{\mathbf{f}}^{(b)}_{T+m|T}$. Then, $\hat{\pi}^{(b)}_{T+m|T}$ is easily obtained by

$$\hat{\pi}^{(b)}_{T+m|T} = RSUM\left(\mathbf{z} \odot \hat{\mathbf{f}}^{(b)}_{T+m|T}\right) \text{ for } b = 1, \ldots, B.$$

Step 4: Robust modifications In practice, forecasting performance of the FAR models may be significantly affected either by the presence of structural breaks or by the fact that inflation and its moments are highly persistent. To address these important issues, we consider two modifications.

Firstly, we follow Clements and Hendry (2006), who identify structural instability as a key factor behind poor forecasting performance, and consider the intercept shift modification by allowing the time-varying mean, $\mathbb{E}[f_t]$ in FAR, where $\mathbb{E}[f_t]$ is estimated by 3-, 6-, 9-, 12-month moving averages of f_t. These modified FAR models, referred to as FAR3M, FAR6M, FAR9M and FAR12M, are given by

$$\text{FARjM} : w_t = A_j w_{t-1} + \epsilon_t,\ j = 3, 6, 9, 12, \tag{A.5}$$

where $w_t = f_t - \bar{f}_{jM}$ and $\bar{f}_{jM} = \frac{1}{j}\sum_{i=0}^{j-1} f_{t-i}$. Similarly, the modified AVE models (AVE3M, AVE6M, AVE9M and AVE12M) are given by

$$\text{AVEjM} : f_t = \bar{f}_{jM} + \epsilon_t,\ j = 3, 6, 9, 12 \tag{A.6}$$

Alternatively, we consider the differenced version of the FAR model, denoted as DFAR, given by

$$w_t = B w_{t-1} + v_t, \tag{A.7}$$

where $w_t = f_t - f_{t-1}$. By construction DFAR removes $\mathbb{E}[f_t]$. Notice that this is a similar strategy as recommended by Clements and Hendry (2002), who suggest to employ the double-differencing of the VAR model for avoiding the forecasting difficulties, see also Kapetanios et al. (2007).

Secondly, the empirical distributions evaluated by the standard bootstrap described in Step 3, will be likely to severely underestimate the realised ones, especially in the presence of extreme events. To address this issue we follow the Generalised Extreme Value Distribution Theory introduced by Embrechts et al. (1997). In particular, Longin (2000) demonstrates that the approach based on extreme values can cover market conditions ranging from the usual environment to extraordinary periods such as the financial crises, especially when outliers may be present in the data. We now propose the robust resampling scheme as follows: We set the block size to h, and collect the non-overlapping sequence of residuals by

$$\left\{\{\hat{\epsilon}_1, \dots, \hat{\epsilon}_h\}, \{\hat{\epsilon}_{h+1}, \dots, \hat{\epsilon}_{2h}\}, \dots, \{\hat{\epsilon}_{T-h+1}, \dots, \hat{\epsilon}_T\}\right\}. \tag{A.8}$$

We then select the functional residuals such that they achieve a maximum for each block as follows:

$$\max_{\left\{\hat{\epsilon}_{h(\tau-1)+i}\right\}_{i=1}^{h}} \left| \int_{-\infty}^{\infty} x \hat{\epsilon}_{h(\tau-1)+i}(x)\, dx \right|,\ \tau = 1, \dots, T(h), \tag{A.9}$$

Then, we draw the b-th bootstrap sample of prediction errors, denoted $\left\{\hat{\epsilon}^{(b)}_{(h)T+1|T}, \dots, \hat{\epsilon}^{(b)}_{(h)T+M|T}\right\}$, from the pool of residuals obtained from (A.9):

$$\left\{\hat{\epsilon}_{(h)1}, \hat{\epsilon}_{(h)2}, \ldots, \hat{\epsilon}_{(h)T(h)}\right\}. \tag{A10}$$

In our empirical analysis we consider $h = 1$ and 3, which are referred to as the standard- and the robust-bootstrap, respectively. There is no optimal rule for selecting h available in the literature, though, theoretically, the selection period should be long enough to guarantee the asymptotic validity of the extreme value theory. In principle, we may conduct the pseudo out-of-sample interval forecasting, and select h by minimising the predictive quantile loss function (e.g., Koenker and Bassett, 1978), as it can be regarded as the predictive quasi-likelihood (e.g., Bertail et al., 2004). Another practically important issue is that too high value of h shrinks the sample size significantly such that no meaningful inference can be made, see also the practical guide discussed by Longin (2000). For our monthly inflation data, the sample size for testing will be only 10 and 20 respectively for $h = 12$ and 6. Hence, our selection of $h = 3$ would make a balance between theory and practice. We note in passing that the results with $h = 2$ or 4 are qualitatively similar to those reported with $h = 3$.

References

Aneiros-Perez, G., Cardot, H., Estevez-Perez, G., and Vieu, P. (2004), "Maximum Ozone Concentration Forecasting by Functional Non-parametric Approaches," *Environmetrics*, 15, 675-685.

Ball, L. and Mankiw, N. G. (1995), "Relative-Price Changes as Aggregate Supply Shocks", *Quarterly Journal of Economics*, 1, 161-193.

Bertail, P., Haefke, C., Politis, D. N. and White, H. (2004), "A Subsampling Approach to Estimating the Distribution of Diverging Statistics with Applications to Assessing Financial Market Risks," *Journal of Econometrics*, 120, 295-326.

Besse, P. C., Cardot, H., and Stephenson, D. B. (2000), "Autoregressive Forecasting of Some Functional Climatic Variations," *Scandinavian Journal of Statistics*, 27, 673-687.

Bils, M. and Klenow, P. (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112, 947-985.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-27.

Bosq, D. (2000), *Linear Processes in Function Spaces*, New York: Springer.

Bowsher, C. G. and Meeks, R. (2008), "The Dynamics of Economic Functions: Modeling and Forecasting the Yield Curve," *Journal of the American Statistical Association*, 103, 1419-1437.

Brier, G. W. (1950), "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review*, 78, 1-3.

Bryan, M. F., Cecchetti, S. G., and Wiggins II, R. L. (1997), "Efficient Inflation Estimation," NBER Working Paper Series, No. 5793.

Cardot, H., Mas, A., and Sarda, P. (1999), "Functional Linear Model," *Statistics & Probability Letters*, 45, 11-22.

Cardot, H., Mas, A., and Sarda, P. (2007), "CLT in Functional Linear Regression Models," *Probability Theory and Related Fields*, 138, 325-361.

Casillas-Olvera, G. and Bessler, D. A. (2006), "Probability Forecasting and Central Bank Accountability," *Journal of Policy Modeling*, 28, 223-234.

Chaudhuri, K., Greenwood-Nimmo, M., Kim, M., and Shin, Y. (2013), "On the Asymmetric U-shaped Relationship between Inflation, Inflation Uncertainty and Relative Price Skewness in the UK," *Journal of Money, Credit and Banking*, 45, 1431-1449.

Chen, S-W., Shen, C-H., and Xie, Z. (2008) "Evidence of a Nonlinear Relationship between Inflation and Inflation Uncertainty: The Case of The Four Little Dragons," *Journal of Policy Modeling* 30, 363-376.

Choi, C. Y. (2010), "Reconsidering The Relationship between Inflation and Relative Price Variability," *Journal of Money, Credit and Banking*, 42, 769-798.

Christoffersen, P. F. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841-861.

Clements, M. P. (2004), "Evaluating The Bank of England Density Forecasts of Inflation," *The Economic Journal*, 114, 844-866.

Clements, M. P. (2010), "Explanations of the Inconsistencies in Survey Respondents' Forecasts," *European Economic Review*, 54, 536-549.

Clements, M. P. and Hendry, D. F. (2002), "Pooling of Forecast," *Econometrics Journal* 5, 1-26.

Clements, M. P. and Hendry, D. F. (2006), "Forecasting with Structural Breaks," in Elliot, G., Granger, C. and Timmermann, A. (edited), *Handbook of Economic Forecasting*, Volume 1, 605-658, Amsterdam: Elsevier.

Cukierman, A. and Meltzer, A. H. (1986), "A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information," *Econometrica* 54, 1099-1128.

Damon, J. and Guillas, S. (2002), "The Inclusion of Exogenous Variables in Functional Autoregressive Ozone Forecasting," *Environmetrics*, 13, 759-774.

Diebold, F. X. and Mariano, R. S. (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253-265.

Diebold, F. X. and Lopez, J. A. (1996), "Forecast Evaluation and Combination," in Maddala, G. S. and Rao, C. R. (edited), *Handbook of Statistics*, Vol. 14: Statistical Methods in Finance, 241-268, Amsterdam: North Holland.

Diebold, F. X., Gunter, T. A., and Tay, A. S. (1998), "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, 863-883.

Elliot, G. and Timmermann, A. (2008), "Economic Forecasting," *Journal of Economic Literature*, 46, 3-56.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997), *Modelling Extreme Events for Insurance and Finance*. Springer, Berlin.

Engelberg, J., Manski, C. F., and Williams, J. (2009), "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters," *Journal of Business and Economic Statistics*, 27, 30-41.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1008.

Faust, J. and Wright, J. H. (2012), "Forecasting Inflation," Mimeo: Johns Hopkins University.

Friedman, M. (1977), "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, 85, 451-472.

Garratt, A., Lee, K., Pesaran, M. H., and Shin, Y. (2006), *Global and National Macroeconometric Modelling: A Long-Run Structural Approach*, Oxford University Press: Oxford.

Giordani, P. and Söderlind, P. (2003), "Inflation Forecast Uncertainty," *European Economic Review*, 47, 1037-1059.

Good, I. J. (1952), "Rational Decisions," *Journal of the Royal Statistical Society, Series B*, 11, 107-114.

Granger, C. W. J. and Pesaran, M. H. (2000), "Economic and Statistical Measure of Forecast Accuracy," *Journal of Forecasting*, 19, 53-70.

Greenwood-Nimmo, M. J., Nguyen, V. H., and Shin, Y. (2012). "Probabilistic Forecasting of Output Growth, Inflation and the Balance of Trade in a GVAR Framework," *Journal of Applied Econometrics*, 27, 554-573.

Hansen, B. E. (1994), "Autoregressive Conditional Density Estimation," *International Economic Review*, 35, 705-730.

Harvey, C. R. and Siddique, A. (2000), "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263-1295.

Holland, A. S. (1995), "Inflation and Uncertainty: Tests for Temporal Ordering," *Journal of Money, Credit and Banking*, 27,827-37.

Hsiao, C., Shen, Y., and Fujiki, H. (2005), "Aggregate vs. Disaggregate Data Analysis: A Paradox in the Estimation of a Money Demand Function of Japan under the Low Interest Rate Policy," *Journal of Applied Econometrics*, 20, 579-602.

Jondeau, E. and Rockinger, M. (2003), "Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements," *Journal of Economic Dynamics and Control*, 27, 1699-1737.

Kapetanios, G., Labhard, V., and Price, S. (2007), "Forecast Combination and The Bank of England's Suite of Statistical Forecasting Models," Working Paper no. 323, Bank of England.

Kargin, V. and Onatski, A. (2008), "Curve Forecasting by Functional Autoregression," *Journal of Multivariate Analysis*, 99, 2508-2526.

Klenow, P. and Malin, B. (2010), "Microeconomic Evidence on Price Setting," NBER Working Paper Series, No. 15826.

Koenker, R. G. and Bassett, G. (1978), "Regression Quantiles," *Econometrica*, 46, 33-50.

Lahiri, K. and Liu, F. (2006), "Modeling Multi-Period Inflation Uncertainty using a Panel of Density Forecasts," *Journal of Applied Econometrics*, 21, 1199-1219.

Laukaitis, A. (2008), "Functional Data Analysis for Cash Flow and Transactions Intensity Continuous-time Prediction Using Hilbert-valued Autoregressive Processes," *European Journal of Operational Research*, 185, 1607-1614.

Longin, F. M. (2000), "From Value at Risk to Stress Testing: The Extreme Value Approach," *Journal of Banking and Finance*, 24, 1097-1130.

Lopez, J. A. (2001), "Evaluating the Predictive Accuracy of Volatility Models," *Journal of Forecasting*, 20, 87-109.

Marzio, M. D. and Taylor, C. C. (2004), "Boosting Kernel Density Estimates: A Bias Reduction Technique?," *Biometrika*, 91, 226–233.

Mas, A. (2007), “Weak Convergence in the Functional Autoregressive Model,” *Journal of Multivariate Analysis*, 98, 1231-1261.

McNeil, A. and Rudiger, F. (2000), “Estimation of Tail-related Risk Measure for Heteroscedastic Financial Time Series: An Extreme Value Approach,” *Journal of Empirical Finance*, 7, 271-300.

Mitchell, J. and Hall, S. G. (2005), “Evaluating, Comparing and Combining Density Forecasts using the KLIC with an Application to the Bank of England and NIESR Fan Charts of Inflation,” *Oxford Bulletin of Economics and Statistics*, 67, 995-1033.

Newey, W. K. and West, K. D. (1987), “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55, 703-708.

Park, J. Y. and Qian, J. (2007), “Autoregressive Modeling of Time-Varying Densities in Functional Space,” Mimeo: Rice University.

Park, J. Y. and Qian, J. (2012), “Functional Regression of Continuous State Distributions,” *Journal of Econometrics*, 167, 397-412.

Pesaran, M. H. and Smith, R. (1995), “Estimating Long-Run Relationships from Dynamic Heterogeneous Panels,” *Journal of Econometrics*, 68, 79-113.

Roger, S. (2000), “Relative Prices, Inflation and Core Inflation,” IMF Working Paper Series, No. 58.

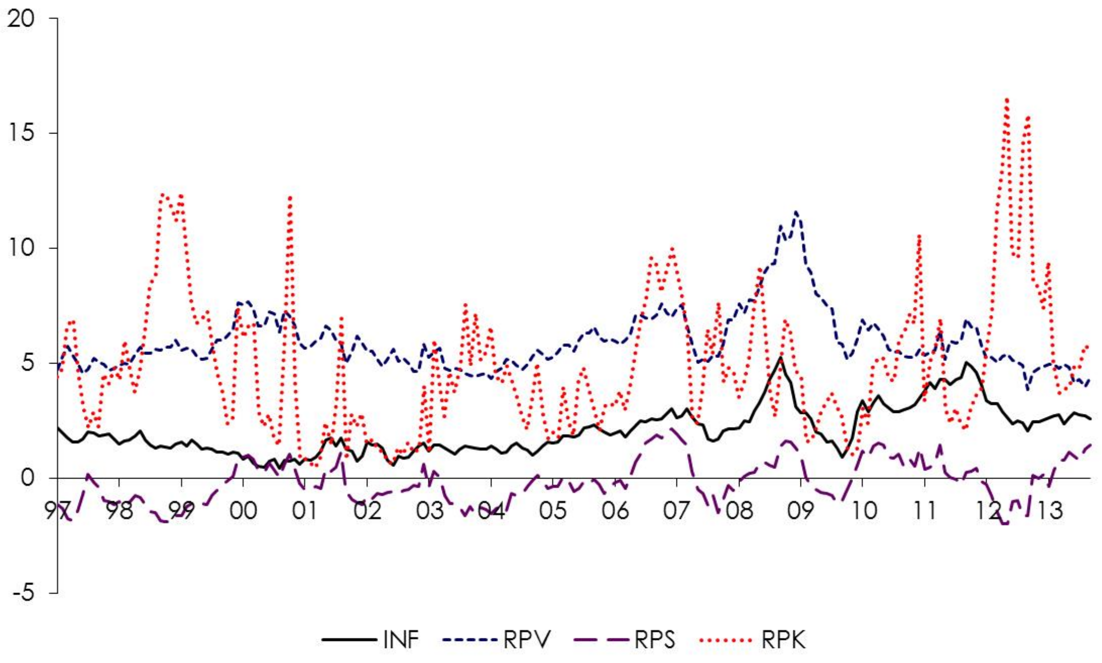
Seillier-Moiseiwitsch, F. and Dawid, A. P. (1993), “On Testing the Validity of Sequential Probability Forecast,” *Journal of American Statistical Association*, 88, 355-359.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall.

Tobin, J. (1972), “Inflation and Unemployment,” *American Economic Review*, 62, 1-18.

Zarnowitz, V. and Llambros, L. A. (1987), “Consensus and Uncertainty in Economic Prediction,” *Journal of Political Economy*, 95, 591-621.

Figure 1: Time-varying moments of sectoral inflation rates.

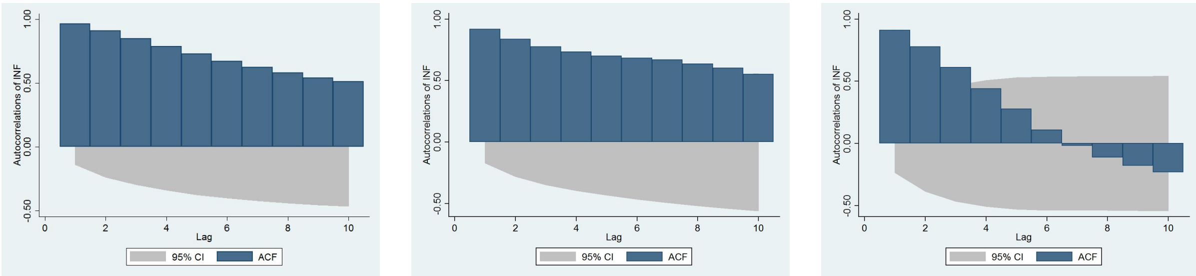


Note: All moments are computed from the estimated cross-sectional density of the sectoral inflation rates by the middle Riemann sum.

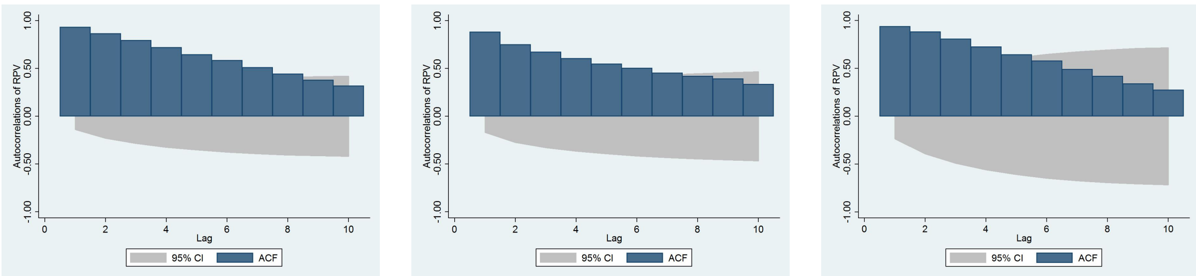
Figure 2: Persistence of cross-sectional moments: autocorrelation function.

(1) Jan. 1997 - Sep. 2013 | (2) Jan. 1997 - Dec. 2007 | (3) Jan. 2008 - Sep. 2013

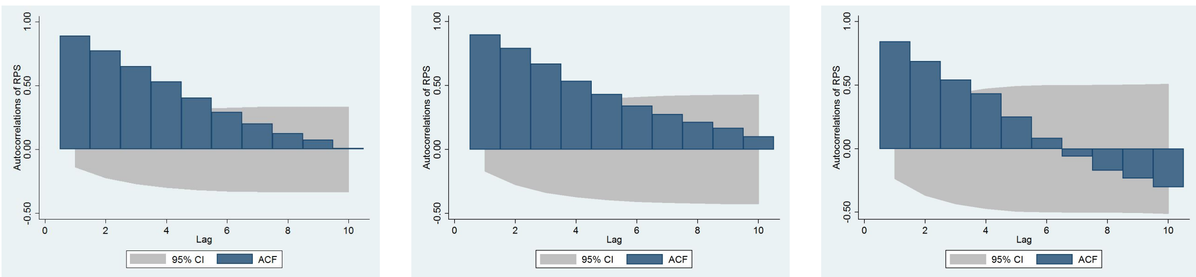
Panel A. INF



Panel B. RPV



Panel C. RPS



Panel D. RPK

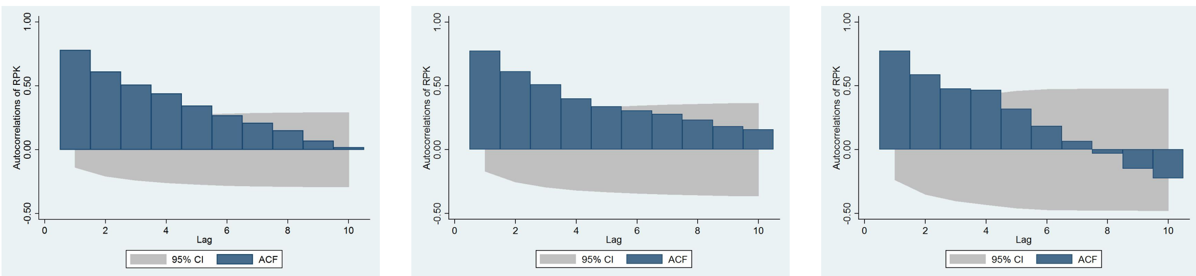


Figure 3: The scatter plots between moments of sectoral inflation rates.

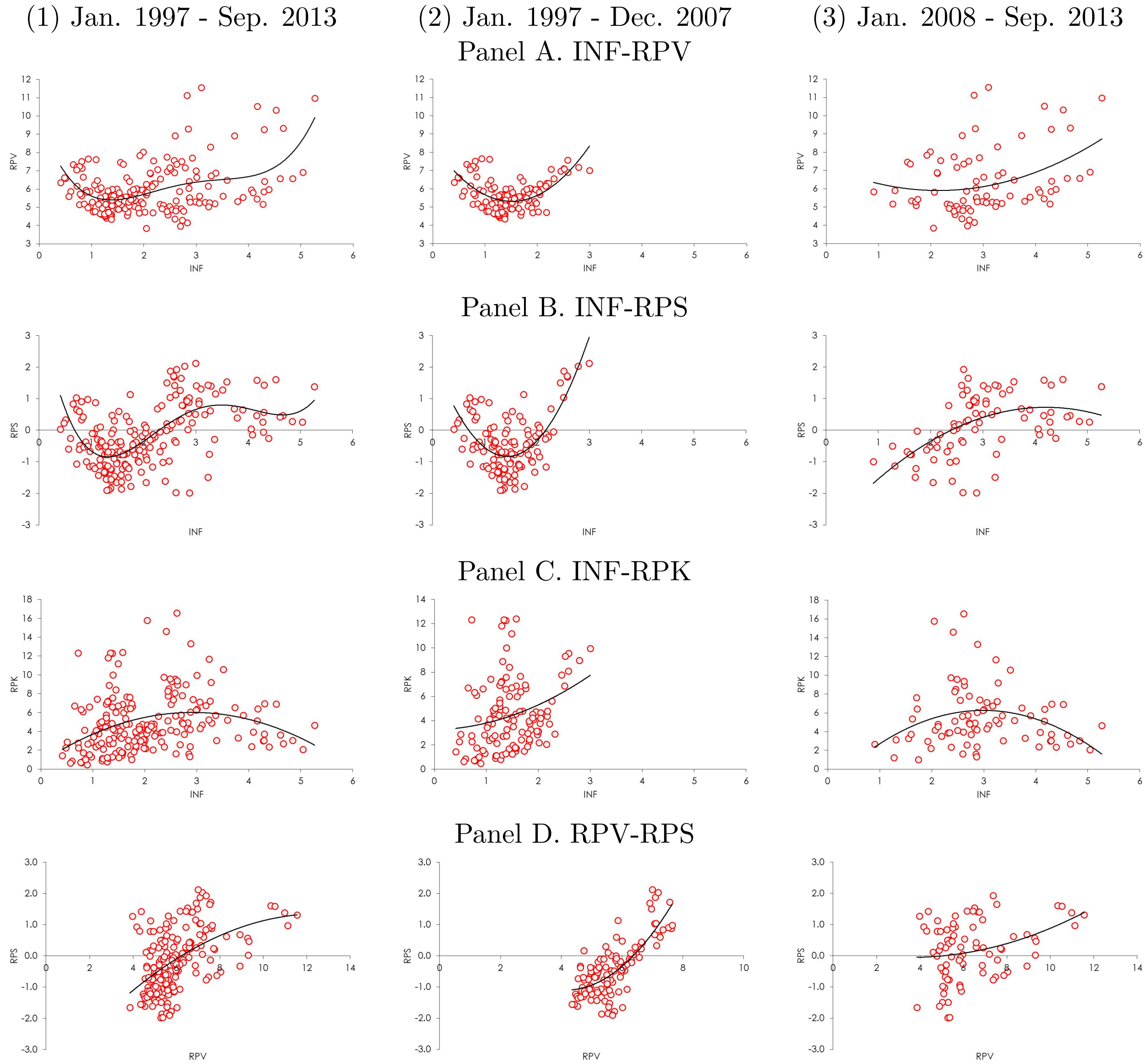


Figure 4: Pseudo out-of-sample interval forecasting of the mean inflation rate: block size = 3-month.

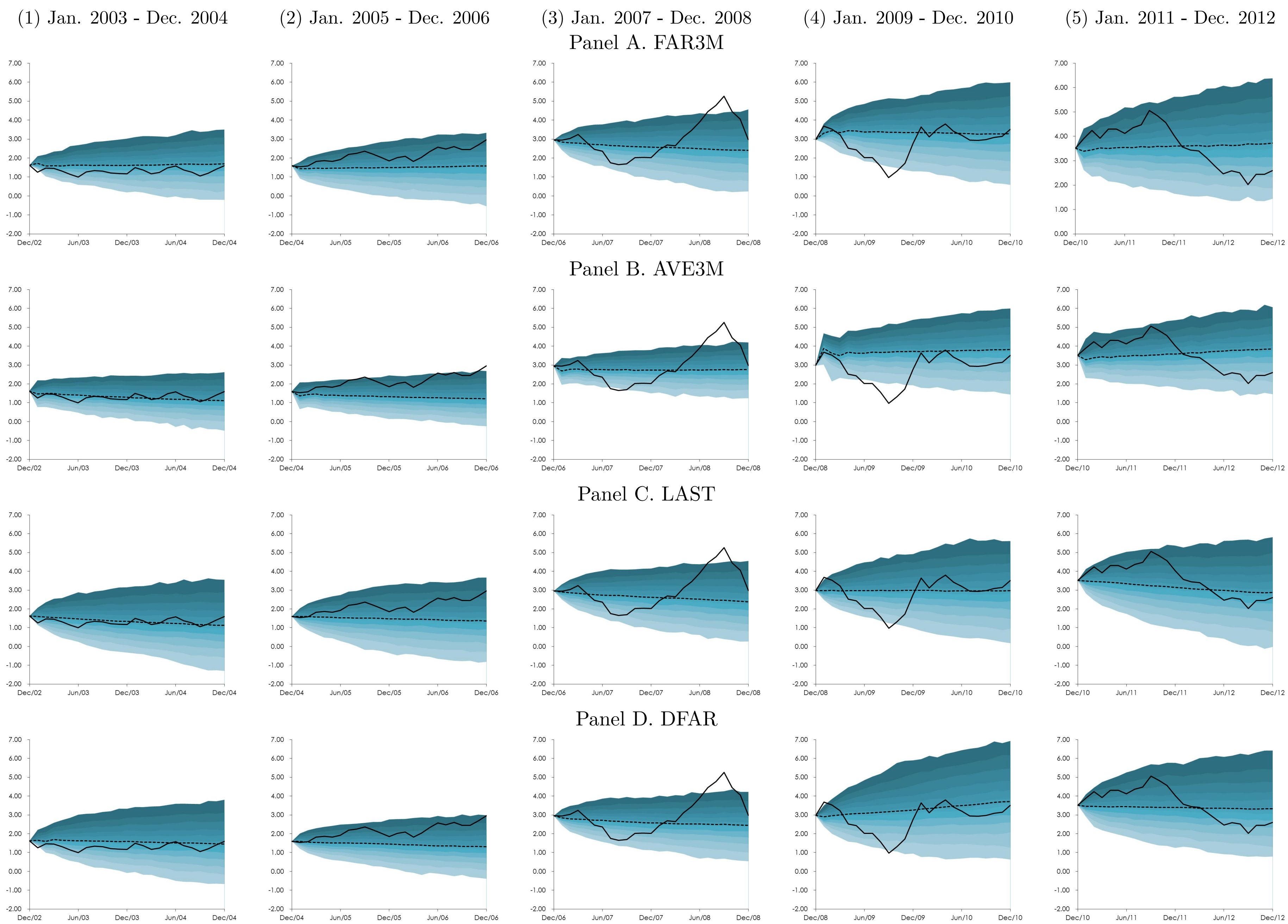
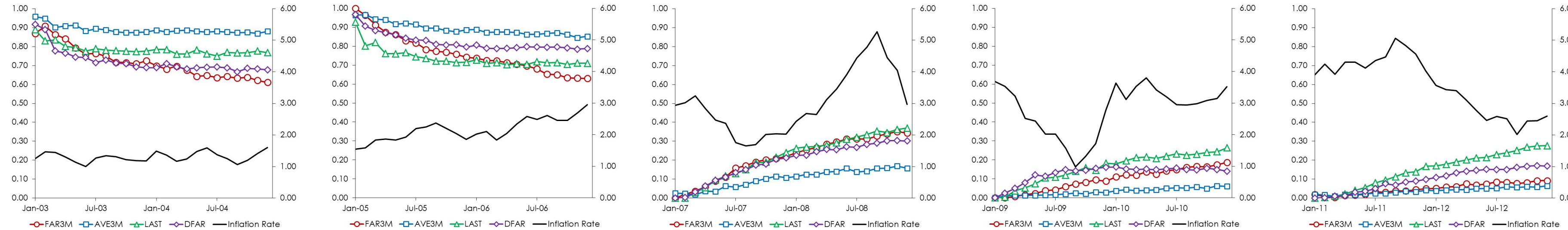


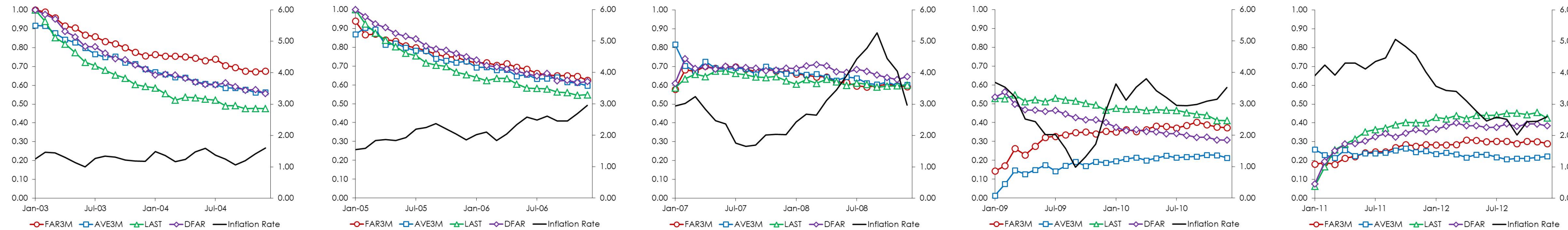
Figure 5: Pseudo out-of-sample probability event forecasting of the mean inflation rate: robust bootstrap: = 3-month.

(1) Jan. 2003 - Dec. 2004 | (2) Jan. 2005 - Dec. 2006 | (3) Jan. 2007 - Dec. 2008 | (4) Jan. 2009 - Dec. 2010 | (5) Jan. 2011 - Dec. 2012

Panel A. $Pr\{\pi < 2\%\}$



Panel B. $Pr\,(1\% < \pi < 3\%)$



Note: Left vertical axis denotes a probability and right vertical axis a realised inflation rate in both panels.

Table 1: Asymmetric relationships between cross-sectional moments.

(1) Jan. 1997 - Sep. 2013			(2) Jan. 1997 - Dec. 2007			(3) Jan. 2008 - Sep. 2013		
β^-	β^+	$H_0: \beta^- = \beta^+$	β^-	β^+	$H_0: \beta^- = \beta^+$	β^-	β^+	$H_0: \beta^- = \beta^+$
Panel A. $INF_t = \left(\alpha^- + \beta^- RPV_t^-\right) + \left(\alpha^+ + \beta^+ RPV_t^+\right) + \epsilon_t$								
-0.106	0.145	11.310	-0.193	0.326	97.810	-0.022	0.085	1.620
(0.052)	(0.053)	[0.000]	(0.027)	(0.045)	[0.000]	(0.044)	(0.072)	[0.208]
Panel B. $INF_t = \left(\alpha^- + \beta^- RPS_t^-\right) + \left(\alpha^+ + \beta^+ RPS_t^+\right) + \epsilon_t$								
-0.140	0.177	12.770	-0.184	0.288	107.630	0.286	0.044	1.690
(0.049)	(0.073)	[0.000]	(0.036)	(0.028)	[0.000]	(0.100)	(0.157)	[0.198]
Panel C. $INF_t = \left(\alpha^- + \beta^- RPK_t^-\right) + \left(\alpha^+ + \beta^+ RPK_t^+\right) + \epsilon_t$								
0.021	-0.055	8.490	0.028	0.058	1.020	0.030	-0.131	21.000
(0.012)	(0.023)	[0.004]	(0.011)	(0.027)	[0.314]	(0.021)	(0.028)	[0.000]
Panel D. $RPV_t = \left(\alpha^- + \beta^- RPS_t^-\right) + \left(\alpha^+ + \beta^+ RPS_t^+\right) + \epsilon_t$								
0.679	0.623	0.100	0.798	0.615	2.210	-0.066	1.401	9.770
(0.092)	(0.148)	[0.749]	(0.112)	(0.051)	[0.139]	(0.265)	(0.387)	[0.003]

Note: We decompose moments into those under the high and the low inflation regimes and apply a threshold regression. INF, RPV, RPS and RPK denote the inflation rate (mean), the relative price variability (standard deviation), the relative price skewness (skewness) and the relative price kurtosis (kurtorsis) of the cross-sectional density of sectoral inflation rates. The high (+) and the low (-) inflation regimes are defined for the case that a mean inflation rate is over and under the overall mean inflation rate: 2.1% (Jan. 1997 - Sep. 2013), 1.5% (Jan. 1997 - Dec. 2007), and 3.1% (Jan. 2008 - Sep. 2013). Figures in $(\cdot)$ and $[\cdot]$ denote the standard errors of coefficients and the p-value of F-test for the null of symmetry, respectively.

Table 2: Cross-sectional density forecasting performance of FAR models

	Mean			Median		
Model	D_H	D_U	D_E	D_H	D_U	D_E
Panel A. Jan. 2003 - Sep. 2013						
FAR	0.0054	0.0971	0.0120	0.0030	0.0784	0.0099
FAR3M	0.0037	0.0821	0.0093	0.0025	0.0714	0.0068
FAR6M	0.0036	0.0804	0.0096	0.0022	0.0731	0.0071
FAR9M	0.0039	0.0844	0.0102	0.0024	0.0744	0.0074
FAR12M	0.0045	0.0897	0.0108	0.0025	0.0767	0.0087
AVE	0.0239	0.1963	0.0359	0.0134	0.1680	0.0271
AVE3M	0.0046	0.0907	0.0110	0.0032	0.0775	0.0080
AVE6M	0.0067	0.1088	0.0152	0.0049	0.1008	0.0112
AVE9M	0.0091	0.1260	0.0189	0.0059	0.1152	0.0163
AVE12M	0.0115	0.1427	0.0226	0.0071	0.1287	0.0184
LAST	0.0034	0.0774	0.0082	0.0023	0.0687	0.0059
DFAR	0.0035	0.0783	0.0085	0.0022	0.0649	0.0060
Panel B. Pre-crisis: Jan. 2003 - Dec. 2007						
FAR	0.0029	0.0764	0.0071	0.0019	0.0689	0.0051
FAR3M	0.0025	0.0717	0.0068	0.0016	0.0653	0.0046
FAR6M	0.0024	0.0690	0.0066	0.0016	0.0640	0.0051
FAR9M	0.0025	0.0723	0.0067	0.0017	0.0677	0.0049
FAR12M	0.0026	0.0733	0.0070	0.0018	0.0709	0.0052
AVE	0.0079	0.1225	0.0158	0.0052	0.1131	0.0164
AVE3M	0.0030	0.0789	0.0078	0.0021	0.0708	0.0057
AVE6M	0.0036	0.0859	0.0097	0.0023	0.0709	0.0071
AVE9M	0.0044	0.0968	0.0115	0.0035	0.0924	0.0093
AVE12M	0.0052	0.1070	0.0133	0.0044	0.1096	0.0115
LAST	0.0022	0.0654	0.0058	0.0015	0.0604	0.0046
DFAR	0.0023	0.0676	0.0059	0.0016	0.0615	0.0040
Panel C. Post-crisis: Jan. 2008 - Sep. 2013						
FAR	0.0076	0.1152	0.0157	0.0042	0.0890	0.0139
FAR3M	0.0047	0.0910	0.0112	0.0033	0.0762	0.0081
FAR6M	0.0047	0.0904	0.0119	0.0031	0.0780	0.0098
FAR9M	0.0051	0.0949	0.0129	0.0035	0.0811	0.0102
FAR12M	0.0060	0.1040	0.0138	0.0039	0.0839	0.0115
AVE	0.0378	0.2605	0.0514	0.0347	0.2495	0.0483
AVE3M	0.0060	0.1009	0.0136	0.0039	0.0811	0.0114
AVE6M	0.0093	0.1287	0.0194	0.0059	0.1088	0.0168
AVE9M	0.0131	0.1514	0.0247	0.0088	0.1344	0.0220
AVE12M	0.0171	0.1737	0.0297	0.0114	0.1516	0.0255
LAST	0.0045	0.0878	0.0100	0.0029	0.0760	0.0073
DFAR	0.0045	0.0875	0.0105	0.0032	0.0766	0.0079

Note: We consider Hilbert norm (DH), uniform norm (DU) and entropy (DE) as divergence criteria, respectively, for mean and median. We estimate models from Jan. 1997 to Dec. 2002, and construct one-month-ahead forecast of cross-sectional density function for Jan. 2003. We repeat this process moving forward one-month at a time in a recursive manner, ending with one-month-ahead density forecast for Sep. 2013 constructed using the estimation results over the period Jan. 1997 – Aug. 2013. This generates a total of 129 observations for evaluating the closeness between one-step-ahead cross-sectional density forecasts and the actual counterparts for each of the twelve models over the forecasting horizon Jan. 2003 - Sep. 2013. Furthermore, we have used 60 and 69 observations in order to evaluate the closeness for the pre-crisis period, Jan. 2003 - Dec. 2007, and the post-crisis period, Jan. 2008 - Sep. 2013.

Table 3: Point forecasting performance of FAR models against AR and ARCH-M: MAE based DM test (Jan. 2003 - Sep. 2013).

Model	Statistic	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M
FAR	MAE	0.406	0.575	0.700	0.783	0.850	0.903	0.944	0.965	0.970	0.979	0.991	1.006
	DM_{AR}	4.671*	1.004	-1.650	2.303*	0.721	0.254	1.814*	0.854	1.059	1.443	0.451	0.321
	DM_{ARCH-M}	-3.954	-3.117	-1.629	-0.287	0.673	1.472	1.829*	2.241*	2.504*	2.613*	2.609*	2.769*
FAR3M	MAE	0.288	0.407	0.512	0.622	0.706	0.788	0.845	0.881	0.899	0.918	0.950	0.980
	DM_{AR}	5.616*	3.326*	1.862*	2.654*	1.696*	1.134	1.413	0.913	0.801	0.846	0.392	0.239
	DM_{ARCH-M}	-1.994	0.887	2.678*	2.967*	3.189*	3.338*	3.413*	3.796*	3.927*	4.045*	3.973*	4.116*
FAR6M	MAE	0.313	0.458	0.568	0.651	0.727	0.800	0.852	0.886	0.914	0.933	0.964	0.992
	DM_{AR}	5.286*	2.442*	1.030	2.411*	1.514*	1.042	1.388*	0.893	0.731	0.778	0.311	0.171
	DM_{ARCH-M}	-2.502	-0.849	0.926	2.139*	2.735*	3.112*	3.291*	3.692*	3.861*	4.006*	3.931*	4.102*
FAR9M	MAE	0.329	0.469	0.586	0.677	0.755	0.809	0.850	0.881	0.912	0.927	0.952	0.969
	DM_{AR}	4.982*	2.353*	0.820	2.357*	1.368	1.022	1.463	0.980	0.793	0.874	0.419	0.335
	DM_{ARCH-M}	-3.241	-1.035	0.422	1.462	2.151*	2.914*	3.320*	3.786*	3.839*	3.985*	4.048*	4.325*
FAR12M	MAE	0.334	0.489	0.601	0.684	0.751	0.796	0.838	0.875	0.898	0.907	0.918	0.926
	DM_{AR}	4.755*	2.040*	0.570	2.298*	1.421*	1.164	1.620*	1.066	0.927	1.050	0.683	0.668
	DM_{ARCH-M}	-3.251	-1.518	0.102	1.144	1.863*	2.643*	3.137*	3.610*	3.860*	4.100*	4.259*	4.620*
AVE	MAE	1.008	1.017	1.027	1.038	1.049	1.060	1.066	1.076	1.087	1.098	1.108	1.116
	DM_{AR}	-3.268	-3.860	-3.748	-1.590	-2.100	-1.912	-0.448	-1.124	-1.184	-0.777	-1.689	-1.779
	DM_{ARCH-M}	-5.129	-4.219	-3.093	-2.088	-1.116	-0.185	0.458	0.987	1.318	1.493	1.620	1.842*
AVE3M	MAE	0.350	0.442	0.539	0.649	0.728	0.789	0.826	0.857	0.883	0.908	0.942	0.950
	DM_{AR}	4.414*	2.651*	1.437*	2.341*	1.464*	1.101	1.535*	1.073	0.911	0.912	0.451	0.426
	DM_{ARCH-M}	-4.258	-0.594	1.966*	2.197*	2.737*	3.325*	3.739*	4.116*	4.168*	4.127*	4.051*	4.463*
AVE6M	MAE	0.485	0.559	0.627	0.691	0.746	0.787	0.831	0.863	0.891	0.913	0.927	0.927
	DM_{AR}	2.269*	0.800	0.170	1.972*	1.285*	1.120	1.526	1.047	0.889	0.921	0.562	0.598
	DM_{ARCH-M}	-5.014	-2.785	-0.370	1.160	2.146*	2.988*	3.239*	3.672*	3.788*	3.839*	3.971*	4.440*
AVE9M	MAE	0.581	0.642	0.693	0.733	0.766	0.796	0.830	0.860	0.883	0.899	0.905	0.905
	DM_{AR}	0.729	-0.368	-0.632	1.554	1.068	1.021	1.561	1.112	0.993	1.071	0.737	0.772
	DM_{ARCH-M}	-4.928	-3.191	-1.185	0.386	1.546	2.596*	3.137*	3.543*	3.667*	3.673*	3.878*	4.347*
AVE12M	MAE	0.626	0.677	0.717	0.750	0.783	0.816	0.846	0.861	0.875	0.879	0.880	0.876
	DM_{AR}	0.131	-0.738	-0.845	1.361	0.901	0.856	1.485	1.125	1.084	1.251	0.958	1.021
	DM_{ARCH-M}	-4.697	-3.098	-1.259	0.140	1.177	2.119*	2.697*	3.310*	3.686*	3.904*	4.088*	4.449*
LAST	MAE	0.256	0.392	0.502	0.609	0.692	0.772	0.822	0.863	0.885	0.906	0.934	0.961
	DM_{AR}	5.805*	3.344*	1.942*	2.693*	1.796*	1.266	1.571	1.042	0.893	0.911	0.491	0.357
	DM_{ARCH-M}	0.614	2.030*	3.353*	3.453*	3.555*	3.752*	3.956*	4.180*	4.221*	4.188*	4.171*	4.389*
DFAR	MAE	0.264	0.383	0.489	0.598	0.681	0.763	0.813	0.860	0.888	0.905	0.933	0.963
	DM_{AR}	5.600*	3.658*	2.184*	2.830*	1.919*	1.354	1.633*	1.067	0.877	0.916	0.501	0.342
	DM_{ARCH-M}	-0.743	2.450*	3.758*	3.694*	3.777*	3.930*	4.056*	4.140*	4.128*	4.174*	4.163*	4.298*
AR	MAE	0.638	0.614	0.640	0.903	0.881	0.913	1.035	1.002	1.015	1.050	1.011	1.019
ARCH-M	MAE	0.258	0.428	0.607	0.764	0.902	1.035	1.132	1.225	1.298	1.353	1.399	1.453

Note: * denotes that DM test is rejected at 5% significance level.

Table 4: Point forecasting performance of FAR models against AR and ARCH-M: MAE based DM test (Jan. 2003 - Dec. 2007).

Model	Statistic	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M
FAR	MAE	0.249	0.358	0.435	0.486	0.515	0.536	0.572	0.591	0.601	0.602	0.619	0.638
	DM_{AR}	4.081*	0.257	-0.684	2.357*	0.529	1.226	3.598*	1.143	2.399*	5.764*	2.418*	4.025*
	DM_{ARCH-M}	-2.122	-1.455	0.065	1.441	2.780*	4.011*	5.103*	6.330*	7.376*	8.124*	9.013*	10.006*
FAR3M	MAE	0.168	0.225	0.253	0.291	0.301	0.315	0.337	0.372	0.381	0.405	0.439	0.474
	DM_{AR}	3.466*	2.494*	2.573*	2.905*	3.143*	3.313*	3.232*	2.822*	2.685*	2.633*	2.357*	2.351*
	DM_{ARCH-M}	-1.199	1.724*	4.165*	6.487*	9.889*	11.697*	11.418*	9.690*	8.244*	7.156*	7.047*	6.979*
FAR6M	MAE	0.162	0.240	0.275	0.292	0.303	0.318	0.341	0.364	0.386	0.415	0.446	0.489
	DM_{AR}	3.815*	2.721*	2.791*	3.147*	3.449*	3.569*	3.399*	3.034*	2.878*	2.844*	2.521*	2.442*
	DM_{ARCH-M}	-0.492	1.744*	5.302*	8.915*	13.272*	14.313*	13.318*	10.849*	9.418*	7.941*	7.747*	7.602*
FAR9M	MAE	0.178	0.255	0.288	0.304	0.313	0.322	0.346	0.372	0.386	0.415	0.452	0.491
	DM_{AR}	3.718*	2.752*	2.843*	3.261*	3.792*	3.642*	3.328*	3.213*	3.315*	3.284*	2.848*	2.737*
	DM_{ARCH-M}	-1.505	1.111	5.328*	9.090*	14.247*	15.061*	13.705*	12.122*	10.944*	9.244*	8.732*	8.467*
FAR12M	MAE	0.192	0.271	0.307	0.331	0.330	0.330	0.355	0.387	0.404	0.423	0.446	0.483
	DM_{AR}	3.684*	2.630*	2.325*	3.063*	3.562*	3.630*	3.455*	3.109*	3.043*	3.148*	2.972*	2.932*
	DM_{ARCH-M}	-2.281	0.314	3.938*	6.636*	10.880*	12.835*	13.190*	11.466*	10.231*	8.774*	8.689*	8.696*
AVE	MAE	0.454	0.487	0.519	0.556	0.587	0.608	0.623	0.631	0.638*	0.645	0.658	0.670
	DM_{AR}	-0.782	-2.073	-2.012	1.081	-1.493	-0.925	3.231*	-0.086	1.577	4.865*	1.073	3.301*
	DM_{ARCH-M}	-2.909	-2.206	-0.885	0.429	1.747*	3.065*	4.393*	5.779*	6.931*	7.703*	8.670*	9.779*
AVE3M	MAE	0.181	0.219	0.243	0.274	0.298	0.317	0.338	0.362	0.383	0.417	0.456	0.489
	DM_{AR}	3.223*	2.675*	2.843*	3.112*	3.247*	3.145*	3.081*	2.958*	2.819*	2.667*	2.184*	2.225*
	DM_{ARCH-M}	-2.107	2.269*	4.962*	8.238*	11.699*	12.176*	11.768*	10.603*	8.844*	7.342*	6.946*	6.995*
AVE6M	MAE	0.202	0.234	0.252	0.277	0.298	0.315	0.341	0.368	0.394	0.430	0.462	0.491
	DM_{AR}	3.565*	3.143*	3.434*	3.321*	3.391*	3.336*	3.271*	3.075*	2.873*	2.681*	2.263*	2.377*
	DM_{ARCH-M}	-1.950	2.363*	6.449*	9.893*	13.306*	13.567*	13.029*	11.268*	9.252*	7.572*	7.320*	7.517*
AVE9M	MAE	0.239	0.265	0.278	0.293	0.309	0.327	0.347	0.377	0.407	0.445	0.473	0.498
	DM_{AR}	3.305*	2.911*	3.166*	3.267*	3.296*	3.267*	3.293*	3.127*	2.928*	2.725*	2.351*	2.499*
	DM_{ARCH-M}	-2.507	0.612	5.932*	9.863*	12.967*	13.763*	13.805*	12.053*	9.974*	8.105*	7.872*	8.173*
AVE12M	MAE	0.256	0.280	0.295	0.315	0.332	0.347	0.368	0.392	0.417	0.448	0.482	0.512
	DM_{AR}	3.113*	2.612*	2.827*	3.144*	3.145*	3.147*	3.240*	3.153*	2.953*	2.764*	2.396*	2.562*
	DM_{ARCH-M}	-2.458	-0.079	4.854*	8.394*	11.610*	13.366*	14.114*	12.872*	10.507*	8.451*	8.319*	8.848*
LAST	MAE	0.152	0.217	0.246	0.279	0.291	0.313	0.341	0.372	0.386	0.400	0.440	0.487
	DM_{AR}	3.427*	2.405*	2.472*	2.849*	3.073*	3.172*	2.971*	2.677*	2.662*	2.703*	2.302*	2.159*
	DM_{ARCH-M}	0.413	1.864*	4.092*	6.495*	9.812*	11.510*	10.745*	9.508*	8.365*	7.210*	6.931*	6.739*
DFAR	MAE	0.156	0.199	0.235	0.271	0.288	0.314	0.335	0.363	0.387	0.400	0.435	0.477
	DM_{AR}	3.269*	2.722*	2.687*	2.934*	3.061*	3.109*	3.039*	2.869	2.682*	2.699*	2.378*	2.283*
	DM_{ARCH-M}	-0.194	2.445*	4.328*	6.497*	9.406*	11.001*	10.552*	9.871	8.253*	7.086*	6.931*	6.787*
AR	MAE	0.421	0.367	0.407	0.581	0.536	0.580	0.676	0.629	0.665	0.700	0.679	0.717
ARCH-M	MAE	0.155	0.278	0.439	0.594	0.736	0.859	0.969	1.055	1.123	1.178	1.238	1.294

Note: * denotes that DM test is rejected at 5% significance level.

Table 5: Point forecasting performance of FAR models against AR and ARCH-M: MAE based DM test (Jan. 2008 - Sep. 2013).

Model	Statistic	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M
FAR	MAE	0.563	0.791	0.970	1.105	1.195	1.257	1.279	1.271	1.227	1.183	1.147	1.123
	DM_{AR}	3.081*	0.746	-1.931	1.258	0.377	-0.304	1.363	0.702	0.710	0.976	0.180	-0.231
	DM_{ARCH-M}	-3.477	-3.018	-2.483	-1.692	-0.948	-0.252	0.053	0.471	0.827	1.033	1.133	1.240
FAR3M	MAE	0.397	0.561	0.726	0.901	1.030	1.153	1.219	1.242	1.251	1.242	1.259	1.286
	DM_{AR}	4.432*	2.633*	1.124	1.769*	1.063	0.444	0.826	0.376	0.126	0.103	-0.362	-0.653
	DM_{ARCH-M}	-1.325	0.122	0.722	0.638	0.753	0.815	0.733	1.247	1.478	1.695*	1.538	1.380
FAR6M	MAE	0.449	0.644	0.815	0.959	1.075	1.181	1.239	1.265	1.284	1.271	1.292	1.305
	DM_{AR}	3.827*	1.736*	0.365	1.458	0.817	0.297	0.740	0.275	-0.010	-0.004	-0.501	-0.739
	DM_{ARCH-M}	-2.225	-1.233	-0.814	-0.275	0.110	0.436	0.472	0.937	1.158	1.411	1.270	1.228
FAR9M	MAE	0.477	0.672	0.857	1.005	1.120	1.194	1.229	1.248	1.284	1.265	1.268	1.260
	DM_{AR}	3.450*	1.527	0.034	1.326	0.629	0.244	0.828	0.369	-0.011	0.019	-0.425	-0.598
	DM_{ARCH-M}	-3.199	-1.441	-1.161	-0.735	-0.388	0.267	0.544	1.025	1.068	1.290	1.317	1.413
FAR12M	MAE	0.482	0.702	0.877	1.007	1.110	1.172	1.209	1.230	1.243	1.214	1.203	1.175
	DM_{AR}	3.264*	1.256	-0.128	1.329	0.705	0.376	0.963	0.470	0.180	0.233	-0.175	-0.283
	DM_{ARCH-M}	-3.088	-1.853	-1.183	-0.579	-0.199	0.352	0.571	0.982	1.230	1.544	1.625	1.784*
AVE	MAE	1.509	1.510	1.508	1.509	1.503	1.490	1.464	1.430	1.395	1.350	1.317	1.289
	DM_{AR}	-4.095	-4.320	-4.173	-1.916	-1.940	-1.836	-0.460	-0.838	-0.997	-0.690	-1.440	-1.830
	DM_{ARCH-M}	-5.676	-4.624	-3.653	-2.796	-1.985	-1.197	-0.701	-0.222	0.111	0.339	0.473	0.607
AVE3M	MAE	0.490	0.626	0.781	0.957	1.069	1.148	1.182	1.203	1.220	1.214	1.232	1.221
	DM_{AR}	3.284*	1.929*	0.660	1.452	0.851	0.466	0.984	0.533	0.248	0.205	-0.260	-0.415
	DM_{ARCH-M}	-3.553	-1.739	-0.484	-0.278	0.204	0.835	1.086	1.518	1.658*	1.753*	1.616	1.727*
AVE6M	MAE	0.730	0.838	0.941	1.039	1.111	1.162	1.208	1.234	1.249	1.237	1.226	1.200
	DM_{AR}	0.923	0.034	-0.614	1.028	0.619	0.389	0.871	0.403	0.138	0.125	-0.251	-0.352
	DM_{ARCH-M}	-5.914	-3.815	-2.360	-1.142	-0.276	0.496	0.592	0.950	1.150	1.318	1.405	1.615
AVE9M	MAE	0.879	0.973	1.054	1.116	1.157	1.183	1.214	1.239	1.245	1.227	1.201	1.171
	DM_{AR}	-0.360	-0.984	-1.353	0.594	0.358	0.270	0.860	0.399	0.162	0.175	-0.157	-0.256
	DM_{ARCH-M}	-5.580	-3.899	-2.555	-1.367	-0.516	0.260	0.506	0.832	1.034	1.161	1.330	1.561
AVE12M	MAE	0.946	1.029	1.097	1.152	1.194	1.229	1.251	1.245	1.229	1.187	1.149	1.113
	DM_{AR}	-0.744	-1.243	-1.497	0.387	0.154	0.040	0.714	0.380	0.238	0.347	0.054	-0.028
	DM_{ARCH-M}	-5.002	-3.562	-2.293	-1.313	-0.647	-0.044	0.219	0.701	1.065	1.363	1.541	1.715*
LAST	MAE	0.355	0.543	0.707	0.877	1.004	1.114	1.167	1.201	1.208	1.213	1.216	1.223
	DM_{AR}	4.713*	2.663*	1.261	1.879*	1.208	0.649	1.065	0.552	0.295	0.207	-0.199	-0.415
	DM_{ARCH-M}	0.474	1.300	1.565	1.305	1.268	1.422	1.443	1.760*	1.924*	1.964*	1.858*	1.812*
DFAR	MAE	0.367	0.536	0.694	0.867	0.990	1.102	1.159	1.208	1.221	1.219	1.225	1.244
	DM_{AR}	4.526*	2.853*	1.397	1.943*	1.286	0.713	1.091	0.522	0.244	0.183	-0.233	-0.485
	DM_{ARCH-M}	-0.605	1.372	1.853*	1.481	1.477	1.585	1.524	1.647*	1.771*	1.917*	1.808*	1.659*
AR	MAE	0.833	0.842	0.861	1.225	1.223	1.236	1.403	1.327	1.282	1.270	1.162	1.106
ARCH-M	MAE	0.358	0.565	0.761	0.939	1.084	1.221	1.288	1.370	1.428	1.455	1.472	1.495

Note: * denotes that DM test is rejected at 5% significance level.

Table 6: Density forecasting performance of FAR models: PIT test

Model	(1) Jan. 2003 - Sep. 2013	(2) Jan. 2003 - Dec. 2007	(3) Jan. 2008 - Dec. 2011	(4) Jan. 2010 - Sep. 2013
Panel A. Block Size = 1-month				
FAR	0.360*	0.366*	0.503*	0.353*
FAR3M	0.149*	0.130	0.266*	0.167
FAR6M	0.143*	0.131	0.298*	0.164
FAR9M	0.172*	0.166	0.352*	0.170
FAR12M	0.169*	0.187*	0.337*	0.141
AVE	0.441*	0.333*	0.706*	0.626*
AVE3M	0.103	0.111	0.253*	0.099
AVE6M	0.134*	0.145	0.300*	0.139
AVE9M	0.176*	0.201*	0.377*	0.173
AVE12M	0.188*	0.242*	0.368*	0.210*
LAST	0.119	0.081	0.289*	0.148
DFAR	0.125*	0.096	0.274*	0.170
Panel B. Block Size = 3-month				
FAR	0.256*	0.269*	0.374*	0.279*
FAR3M	0.139*	0.243*	0.200*	0.154
FAR6M	0.098	0.187*	0.216*	0.091
FAR9M	0.094	0.139	0.275*	0.121
FAR12M	0.101	0.133	0.276*	0.124
AVE	0.399*	0.391*	0.677*	0.619*
AVE3M	0.102	0.177*	0.165	0.170
AVE6M	0.116	0.218*	0.255*	0.108
AVE9M	0.167*	0.286*	0.336*	0.174
AVE12M	0.141*	0.252*	0.307*	0.161
LAST	0.143*	0.198*	0.254*	0.144
DFAR	0.096	0.116	0.195	0.107

Note: Kolmogorov-Smirnov (KS) statistics is adopted for testing the equality of empirical distribution of PITs with uniform distribution. The K-S test statistics is obtained by $\max\{D_+, D_-\}$, where $D_+ = \max_x \left\{\hat{F}(x) - U(x)\right\}$ and $D_- = \max_x \left\{U(x) - \hat{F}(x)\right\}$. $\hat{F}(x)$ is the empirical cumulative probability function of PIT and $U(x)$ is the uniform cumulative probability distribution on the closed interval $[0, 1]$. The PIT test is evaluated using 129 PITs for the forecasting period of Jan. 2003 - Sep. 2013, 60 PITs for Jan. 2003 - Dec. 2007, 69 PITs for Jan. 2008 - Sep. 2013, 48 PITs for Jan. 2008 - Dec. 2011, and 45 PITs for Jan. 2010 - Sep. 2013. * denotes that the the KS test is rejected at the 5% significance level. We use a 3-month as the block size for the robust bootstrap.

Table 7: Interval forecasting performance of FAR models: CC test for 95% forecasting interval

Model	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M
Panel A. Block Size = 1-month												
1. Pre-crisis: Jan. 2003 - Dec. 2007												
FAR3M	5.12	0.25	2.41	1.13	0.25	1.13	3.36	1.48	16.89*	11.46*	6.66*	6.66*
AVE3M	1.13	1.13	1.13	0.25	0.55	5.88	11.41*	22.19*	33.13*	20.14*	30.61*	25.31*
LAST	4.92	4.92	1.13	4.92	1.13	4.92	1.09	1.09	0.41	9.18*	6.66*	0.25
DFAR	0.34	1.13	1.13	4.92	4.92	4.92	1.09	0.16	0.16	1.13	2.41	1.13
2. Post-crisis: Jan. 2008 - Dec. 2011												
FAR3M	35.90*	45.66*	44.31*	65.40*	81.02*	69.24*	69.24*	68.51*	63.58*	35.11*	30.37*	26.40*
AVE3M	35.90*	58.03*	85.21*	97.95*	99.30*	86.01*	75.09*	80.12*	71.39*	55.55*	42.89*	42.95*
LAST	20.08*	32.33*	52.31*	68.51*	75.09*	68.51*	52.14*	51.52*	41.21*	38.21*	19.92*	21.98*
DFAR	30.73*	33.50*	33.50*	46.59*	68.51*	62.79*	46.59*	51.52*	41.21*	25.09*	21.95*	21.98*
3. Post-crisis: Jan. 2010 - Sep. 2013												
FAR3M	1.76	2.00	6.35*	6.35*	6.35*	11.69*	7.03*	7.03*	19.20*	19.20*	24.76*	27.56*
AVE3M	0.44	7.03*	6.35*	20.05*	30.33*	38.68*	33.08*	32.25*	38.68*	33.08*	43.33*	44.33*
LAST	0.44	0.37	7.03*	8.46*	16.45*	11.69*	11.69*	24.76*	19.20*	24.76*	16.82*	22.13*
DFAR	6.84*	3.39	2.54	2.54	11.69*	11.69*	11.69*	24.76*	19.20*	16.82*	24.76*	22.13*
Panel B. Block Size = 3-month												
1. Pre-crisis: Jan. 2003 - Dec. 2007												
FAR3M	1.18	1.13	4.92	4.92	4.92	4.92	4.92	1.13	4.92	4.92	1.13	4.92
AVE3M	4.92	1.13	4.92	4.92	4.92	4.92	1.09	0.41	6.92*	11.46*	11.46*	2.41
LAST	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92
DFAR	1.18	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92	4.92
2. Post-crisis: Jan. 2008 - Dec. 2011												
FAR3M	9.60*	33.50*	38.84*	49.88*	42.89*	40.50*	37.54*	46.72*	29.82*	32.72*	21.95*	21.98*
AVE3M	23.26*	30.73*	55.55*	62.79*	63.46*	62.79*	57.12*	63.58*	54.18*	30.37*	29.82*	26.40*
LAST	12.71*	21.03*	23.26*	26.80*	32.33*	45.97*	46.59*	35.11*	41.21*	14.76*	14.91*	27.40*
DFAR	21.03*	16.63*	26.80*	32.07*	32.33*	40.50*	29.82*	35.11*	35.75*	21.95*	16.73*	19.30*
3. Post-crisis: Jan. 2010 - Sep. 2013												
FAR3M	0.36	3.52	2.57	2.57	6.88*	3.52	2.57	8.43*	13.71*	13.71*	8.43*	11.70*
AVE3M	3.28	0.34	3.52	2.57	16.88*	11.74*	11.74*	19.17*	16.88*	16.88*	27.63*	30.37*
LAST	0.34	3.28	0.34	3.28	3.28	0.34	3.52	8.43*	0.34	1.66	2.57	2.57
DFAR	3.28	3.28	3.28	0.34	3.28	3.28	0.34	3.28	0.34	0.36	0.34	0.34

Note. We evaluate the coverage ability of model by the conditional coverage test developed by Christofferesen (1998). * denotes that the test is rejected at the 5% significance level.

Table 8: Probability event forecasting of FAR models: QPS

Model	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M
Panel A. $\pi < 2\%$												
1. Full period: 2003–2013												
FAR3M	0.130	0.176	0.222	0.265	0.290	0.312	0.325	0.332	0.328	**0.331**	0.350	**0.359**
AVE3M	0.151	0.198	0.238	0.274	0.295	0.312	**0.317**	**0.324**	**0.327**	0.336	**0.349**	0.360
LAST	0.126	0.174	0.226	0.260	0.294	0.320	0.335	0.347	0.349	0.362	0.379	0.397
DFAR	**0.116**	**0.170**	**0.213**	**0.247**	**0.279**	**0.303**	0.319	0.345	0.355	0.360	0.386	0.405
2. Pre-crisis: 2003 – 2007												
FAR3M	0.085	0.122	0.146	**0.181**	**0.201**	**0.209**	**0.237**	**0.276**	**0.305**	**0.330**	**0.366**	**0.365**
AVE3M	0.098	**0.120**	**0.140**	0.183	0.221	0.253	0.273	0.312	0.337	0.369	0.410	0.443
LAST	0.086	0.131	0.150	0.185	0.224	0.249	0.276	0.310	0.330	0.360	0.386	0.410
DFAR	**0.070**	0.133	0.148	0.190	0.229	0.248	0.265	0.309	0.349	0.366	0.424	0.450
3. Crisis & Post-crisis: 2008–2013												
FAR3M	0.071	**0.101**	**0.150**	0.206	0.244	0.283	0.298	0.305	0.300	0.295	0.308	0.314
AVE3M	0.089	0.145	0.196	0.237	0.253	0.271	**0.276**	**0.276**	**0.275**	**0.275**	**0.272**	**0.274**
LAST	0.069	0.104	0.160	0.195	0.232	0.268	0.291	0.306	0.314	0.329	0.339	0.356
DFAR	**0.053**	0.102	0.155	**0.186**	**0.218**	**0.254**	0.284	0.316	0.327	0.335	0.346	0.361
Panel B. $1\% < \pi < 3\%$												
1. Full period: 2003–2013												
FAR3M	0.179	0.253	0.301	0.331	0.358	0.371	0.385	0.434	0.464	**0.477**	**0.498**	**0.507**
AVE3M	0.207	0.250	0.306	0.355	0.384	0.416	0.432	0.468	0.501	0.526	0.537	0.540
LAST	0.166	0.223	0.286	0.310	0.348	0.370	0.392	0.433	0.476	0.502	0.519	0.523
DFAR	**0.158**	**0.220**	**0.263**	**0.285**	**0.321**	**0.333**	**0.355**	**0.406**	**0.452**	0.483	0.510	0.518
2. Pre-crisis: 2003 – 2007												
FAR3M	0.032	0.050	0.057	0.067	0.078	0.087	0.097	0.106	**0.114**	0.126	**0.138**	**0.148**
AVE3M	0.060	0.069	0.083	0.099	0.115	0.130	0.148	0.163	0.176	0.192	0.207	0.222
LAST	0.051	0.074	0.089	0.107	0.122	0.143	0.157	0.174	0.193	0.206	0.226	0.247
DFAR	**0.031**	**0.035**	**0.049**	**0.056**	**0.067**	**0.078**	**0.093**	**0.105**	0.117	**0.125**	0.140	0.155
3. Crisis & Post-crisis: 2008–2013												
FAR3M	0.331	0.464	0.544	0.586	0.587	0.584	0.581	0.654	0.685	**0.685**	**0.710**	0.730
AVE3M	0.368	0.445	0.535	0.616	0.633	0.654	0.638	0.667	0.689	0.712	0.718	0.721
LAST	0.275	**0.381**	0.499	0.517	0.554	0.564	0.578	0.628	0.686	0.714	0.725	0.715
DFAR	**0.274**	0.397	**0.473**	**0.502**	**0.536**	**0.535**	**0.545**	**0.609**	**0.672**	0.700	0.721	**0.710**

Note. We evaluate the probability event forecasting of model by QPS by Brier (1950). Bold font indicates the smallest QPS among four functional models.

Table 9: Probability event forecasting of FAR models: Calibration test

MODEL	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M
Panel A. $\pi < 2\%$												
1. Full period: 2003–2013												
FAR3M	-1.21	-1.84	-2.07*	-2.19*	-2.30*	-2.37*	-2.47*	-2.77*	-2.86*	-3.16*	-3.43*	-3.33*
AVE3M	-1.39	-1.62	-1.89	-2.32*	-2.55*	-2.76*	-2.96*	-3.18*	-3.31*	-3.75*	-4.02*	-4.38*
LAST	-1.46	-1.59	-1.93	-2.16*	-2.22*	-2.55*	-2.89*	-3.02*	-3.19*	-3.49*	-3.80*	-4.15*
DFAR	-1.50	-1.63	-1.77	-1.92	-2.20*	-2.44*	-2.63*	-2.79*	-3.16*	-3.38*	-3.57*	-3.84*
2. Pre-crisis: 2003 – 2007												
FAR3M	0.84	0.19	0.42	0.21	-0.02	-0.15	-0.31	-0.63	-0.85	-1.25	-1.57	-1.45
AVE3M	0.24	-0.13	-0.13	-0.72	-1.24	-1.63	-2.02*	-2.51*	-2.88*	-3.42*	-3.94*	-4.53*
LAST	0.79	0.22	0.40	0.29	0.15	-0.15	-0.59	-0.61	-0.87	-1.17	-1.51	-1.83
DFAR	0.48	0.13	0.50	0.38	0.02	-0.23	-0.47	-0.70	-0.97	-1.27	-1.56	-1.86
3. Crisis & Post-crisis: 2008–2013												
FAR3M	-1.93	-2.08*	-2.11*	-2.01*	-2.05*	-2.03*	-1.96*	-2.09*	-1.98*	-2.05*	-2.12*	-2.07*
AVE3M	-1.56	-1.59	-1.37	-1.47	-1.30	-1.30	-1.24	-1.18	-1.20	-1.30	-1.18	-1.23
LAST	-1.90	-1.69	-1.78	-1.98*	-2.02*	-2.07*	-2.16*	-2.25*	-2.33*	-2.42*	-2.53*	-2.66*
DFAR	-1.78	-1.95	-2.02*	-2.08*	-2.11*	-2.23*	-2.23*	-2.29*	-2.44*	-2.56*	-2.59*	-2.71*
Panel B. $1\% < \pi < 3\%$												
1. Full period: 2003–2013												
FAR3M	1.64	1.59	1.52	1.53	1.46	1.42	1.77	1.98*	2.15*	2.09*	2.17*	2.11*
AVE3M	2.03*	1.75	1.75	1.81	1.67	1.59	1.87	2.01*	2.14*	2.04*	2.05*	1.99*
LAST	1.63	1.81	1.86	1.75	1.79	2.06*	2.11*	2.41*	2.69*	2.68*	2.61*	2.52*
DFAR	1.44	1.15	1.25	1.22	1.13	1.17	1.49	1.75	1.94	1.80	1.80	1.79
2. Pre-crisis: 2003 – 2007												
FAR3M	1.99*	2.33*	2.49*	2.67*	2.86*	3.04*	3.13*	3.23*	3.28*	3.39*	3.53*	3.69*
AVE3M	2.56*	2.67*	2.87*	3.18*	3.29*	3.40*	3.54*	3.65*	3.78*	4.00*	4.13*	4.28*
LAST	2.34*	2.82*	3.01*	3.08*	3.28*	3.68*	3.71*	3.88*	4.01*	4.15*	4.37*	4.43*
DFAR	2.06*	2.17*	2.39*	2.61*	2.71*	2.84*	3.07*	3.24*	3.32*	3.39*	3.55*	3.78*
3. Crisis & Post-crisis: 2008–2013												
FAR3M	-0.94	-1.44	-1.94	-2.17*	-2.19*	-2.27*	-1.98*	-1.50	-1.02	-0.72	-0.28	0.18
AVE3M	-0.95	-1.32	-1.62	-1.99*	-1.86	-1.86	-1.64	-1.19	-0.76	-0.45	0.04	0.41
LAST	-1.15	-1.49	-1.99*	-2.37*	-2.20*	-2.14*	-2.23*	-1.89	-1.56	-1.14	-0.87	-0.46
DFAR	-0.88	-1.59	-1.81	-2.26*	-2.29*	-2.06*	-2.13*	-1.82	-1.32	-0.89	-0.51	-0.13

Note. We evaluate the probability event forecasting of model using the calibration test by Diebold and Lopez (1996). * denotes that the test is rejected at the 5% significance level.