Are Macroeconomic Variables Useful for Forecasting the Distribution of U.S. Inflation?

Sebastiano Manzan[a] and Dawit Zerom[b]

[a] *Department of Economics & Finance, Baruch College, CUNY*

[b] *California State University at Fullerton*

Abstract

Much of the inflation forecasting literature examines the ability of macroeconomic indicators to accurately predict mean inflation. For the period after 1984, existing empirical evidence largely suggests that the likelihood of accurately predicting inflation using macroeconomic indicators is no better than a random walk model. We expand the scope of inflation predictability by exploring whether macroeconomic indicators are useful in predicting the distribution of inflation. We consider six commonly used macro indicators and core/non-core versions of the Consumer Price Index (CPI) and the Personal Consumption Expenditure (PCE) deflator as measures of inflation. Based on monthly data and for the forecast period after 1984, we find that some of the macro indicators, such as unemployment rate and housing starts, provide significant out-of-sample predictability for the distribution of future core inflation. The analysis of the quantiles of the predictive distribution reveals interesting patterns which otherwise would be ignored by existing inflation forecasting approaches that rely only on forecasting the mean. We also illustrate the importance of inflation distribution forecasting in evaluating some events of policy interest by focusing on predicting the likelihood of deflation.

JEL Classification: C22, C53, E31, E52

Keywords: Distribution; Economic indicators; Inflation; Predictability; Quantiles.

Acknowledgements: We are grateful to Barbara Rossi and seminar participants at Baruch College, Rutgers University, SNDE 2009 Symposium, and the 2009 NBER Summer Institute for helpful comments that significantly improved the paper.

1 Introduction

Forecasting the behavior of inflation plays a central role in the conduct of monetary policy due to the lagged impact of the central bank actions on economic activity. It is thus important to accurately predict the effect of the many shocks that hit the economy on the future dynamics of inflation. The standard approach for forecasting inflation has been the Phillips curve (PC) model that, in its expectation-augmented version, assumes a trade-off between unexpected inflation and unemployment, or more generally, indicators of real economic activity. Despite its long-time success, recent empirical evidence on the effectiveness of the PC model is far from unanimous. Stock and Watson (1999) provide a detailed study on the out-of-sample forecast accuracy of the PC by using an extensive set of macroeconomic variables. Using the forecast evaluation period January 1970 - September 1996, their conclusion is that PC models have better forecasting performances (compared to univariate time series models) using the unemployment rate as well as other leading indicators of economic activity (e.g., output gap and capacity utilization). They also find that combining information or models might provide better results than simply relying on few indicators. However, Atkenson and Ohanian (2001) provide an opposite empirical evidence, albeit a different forecast evaluation period January 1984 - November 1999, where they report that PC models are no better than the naïve model, which assumes that the expected inflation over the next 12 months is equal to inflation over the previous 12 months. For a comprehensive survey as well as discussion of the outstanding issues in inflation forecasting, see Stock and Watson (2008).

While the contrasting findings on inflation predictability cannot be directly compared, they may, nevertheless, indicate that the relationship predicted by the PC models might have been unstable over time due to a possible shift in the dynamics of inflation. A phenomenon that is typically suggested to have caused a regime shift is the change in monetary policy that took place when Paul Volcker became Chairman of the Federal Reserve Board in August 1979. The effect of the stricter monetary policy was fully incorporated into the inflation process after 1984 and since then inflation has been low and stable (compared to the 1970s). Fisher *et al.* (2002) conduct a systematic comparison of the forecasting accuracy (one-year ahead) of the naïve and PC models in different sub-periods: January 1977 to December 1984, January 1985 to December 1992 and January 1993 to December 2000. They find that the PC forecasts outperform the naïve forecasts only in the first sub-period for most of the inflation measures that they consider. The issue of model instability is also examined by Clark and McCracken (2006) and their results suggest that while model instability cannot be ruled out, the bulk of the findings of unpredictability could also be the result of the low power of out-of-sample forecast comparison tests. Inoue and Kilian (2004) also question the practice of evaluating a model based on its out-of-sample performance (as opposed to its in-sample fit).

Notwithstanding the causes (e.g., regime change), most of the current empirical evidence suggests that indicators of economic activity are weak predictors of inflation. This is especially true in the most recent years (post 1984) when forecasting inflation has become increasingly harder in the sense of providing forecast gains over time series models (Stock and Watson, 2007). Despite the availability of extensive literature on inflation forecasting, little or no attention has been paid to examining whether indicators of economic activity carry useful information about the dynamics of higher moments, beyond the mean. For example, having some idea on the conditional second-order moment of future inflation can be vital in assessing the risk to inflation stability due to macroeconomic shocks. Greenspan (2004) discusses this issue in the following terms: "Given our inevitably incomplete knowledge about key structural aspects of an ever-changing economy and the sometimes asymmetric costs or benefits of particular outcomes, a central bank needs to consider not only the most likely future path for the economy, but also the distribution of possible outcomes about that path. The decision-makers then need to reach a judgment about the probabilities, costs, and benefits of the various possible outcomes under alternative choices for policy" (p. 37). While average future inflation may signal the direction of the economy, it cannot help policy-makers to evaluate the risks of deviations from the most likely path and the cost for the economy of such deviations. In a recent paper, Kilian and Manganelli (2008) introduce a model in which the monetary policy maker is viewed as a risk manager trying to balance the risks to inflation and output stability. In this framework, if the preferences of the policy maker are assumed to be quadratic and symmetric, then the only relevant moment (of the inflation and output distributions) is the conditional mean. However, they provide evidence of departure of the preferences from such a benchmark. All the above elements point to the suggestion that forecasting the distribution of inflation represents a relevant tool in the conduct of monetary policy. In fact, the Bank of England has been publishing for many years the so-called "fan charts" that represent the subjective forecasts of the Bank about the future distribution of inflation.

In this paper, departing from the existing focus on conditional mean forecasting, we explore whether leading indicators of economic activity are useful in predicting the distribution of future inflation. We use linear quantile regression to incorporate macroeconomic variables into the prediction of the conditional distribution of future inflation. The approach considers several conditional quantiles of future inflation, and by doing so, offers more flexibility (than, for example, the conventional PC models) in capturing the possible role of macroeconomic indicators in predicting the different parts of the inflation distribution. For instance, one may be able to investigate if some periods of low or high inflation are driven by some macroeconomic indicators. Surely such information cannot be delivered by PC-type models that deal only with predicting average inflation.

We find strong empirical evidence of predictability of U.S. core monthly inflation, for which we find that indicators of economic activity are useful in forecasting its distribution, especially when using unemployment rate and housing starts. Importantly, the empirical findings apply to a forecast evaluation period that is intentionally chosen to be post 1984, when the existing literature shows that macroeconomic indicators are not relevant to predict future average inflation. To the best of our knowledge this empirical evidence is new in the US inflation forecasting literature. We attribute this result to the ability of our approach to account for the varying predictive effect of economic indicators on core inflation at different quantiles of its distribution. For some indicator variables, we also find an asymmetric effect in the sense that an indicator is more relevant on the lower part of the forecasting distribution than the upper part (and vice versa, depending on the indicator considered). These observed quantile effects take place far away from the center of the distribution, making them difficult to be detected with approaches (like PC-type models) that solely focus on evaluating the relevance of these variables in predicting the conditional mean.

A few, yet increasing, research exists that relates to our work in the sense of dealing with distributional aspects of inflation. Robertson *et al.* (2005) forecast the distribution of inflation based on a VAR specification. In addition, they propose a methodology to "twist" the forecasting distribution in order to incorporate theoretical restrictions (e.g., a Taylor rule). Cogley *et al.* (2005) propose a Bayesian VAR model where both the conditional mean and variance are time varying. They forecast inflation for the UK and illustrate their method by comparing interval forecasts from their model to the fan charts of the Bank of England. Corradi and Swanson (2006) evaluate the performance of time series and PC models in forecasting one-month ahead inflation using different distributional assumptions for the error term. Amisano and Giacomini (2007) forecast the distribution of inflation at the one-month horizon using a Markov Switching model and find that the forecasts from the nonlinear specification are more accurate compared to a linear one.

The rest of the paper is organized as follows. In Section (2) we discuss econometric approaches, including our proposal, to estimate the conditional forecast distribution of future inflation. The proposed idea is very simple and builds on existing well known methods. Section (3) outlines a test of predictive accuracy that is used to evaluate the conditional distribution models discussed in Section (2). In Section (4), we present (with discussion) the empirical findings of the paper. Finally, Section (5) concludes.

2 Econometric methodology

We denote the annualized inflation over a h-month period by $Y_t^h = (1200/h)[\log P_t - \log P_{t-h}]$ and the one-month annualized inflation by $Y_t = 1200[\log P_t - \log P_{t-1}]$ where P_t is the level of the price index in month t. Also let X_t^i be some indicator of real economic activity such as unemployment rate. A baseline specification often used in forecasting inflation is the Phillips curve (PC) model. The PC model (see Stock and Watson, 1999) postulates that changes in h-month inflation, Y_{t+h}^h, depend on recent changes in one-month inflation and past and present values of a candidate economic indicator,

$$Y_{t+h}^h - Y_t = \mu_0^h + \beta^h(L)\Delta\ Y_t + \gamma^h(L)X_t^i + U_{t+h}^{pc} \tag{1}$$

where μ_0^h is a constant, $\beta^h(L)$ and $\gamma^h(L)$ are lag polynomials written in terms of the lag operator L, U_{t+h}^{pc} is the h-step ahead error term with zero conditional mean. Note that the above specification assumes that Y_t has a unit root. There is no consensus yet on the stationarity of inflation (see the recent work by Stock and Watson, 2007, and Ang *et al.*, 2006, for opposite views). In evaluating the forecasting performance of the PC model, it is often compared against two univariate models: the autoregressive (AR) model and the naïve (random walk) model. Although simple, these two time series models are very competitive benchmarks. The naïve model (see Atkenson and Ohanian, 2001) specifies that the expected inflation over the next h months is equal to inflation over the previous h months, i.e.,

$$Y_{t+h}^h - Y_t^h = U_{t+h}^{ao} \tag{2}$$

where the error term U_{t+h}^{ao} has a zero conditional mean. In the rest of the paper we will refer to the naïve model by AO model. The AR model is a special case of the PC model where no information on present and past values of X_t^i are included, i.e.,

$$Y_{t+h}^h - Y_t = \mu_0^h + \beta^h(L)\Delta\ Y_t + U_{t+h}^{ar} \tag{3}$$

where the error term U^{ar}_{t+h} has a zero conditional mean.

2.1 Proposed approach

Note from the PC model specification that its error U^{pc}_{t+h} is assumed to be independent of the past and present values of the economic indicator (X^i_t). In other words, the effect (if any) of the macroeconomic variable on the conditional distribution of future inflation (Y^h_{t+h}) is only limited to the conditional mean. So the PC approach ignores the possibility for economic indicators to carry useful information about the dynamics of higher moments, and hence help improve the accuracy of the forecast distribution of inflation. To circumvent this limitation of PC models, we focus instead on the forecast distribution of Y^h_{t+h} conditional on the available information set at time t. Let the past and present values of a particular economic indicator variable, X^i, be denoted by the vector $\tilde{X}^i_t = (X^i_t, \cdots, X^i_{t-q+1})$. Our approach assumes the naïve model (see Equation 2) to be the true model for predicting the conditional mean of Y^h_{t+h}. This implies that $\tilde{X}_t$ does not carry any relevant information for predicting the mean future inflation. The choice of the naïve model for the conditional mean of the inflation process is motivated by the overwhelming evidence that this specification outperforms PC models in out-of-sample forecasting (at least for the post 1984 period). However, unlike the AO model, we allow $\tilde{X}^i_t$ to have an effect on higher-order moments of Y^h_{t+h} and, more generally, on the conditional distribution of Y^h_{t+h}, which is not permitted in the PC model. Under this set-up, if we find (relative) predictability in the conditional distribution of Y^h_{t+h} using $\tilde{X}^i_t$ while no (relative) predictability by the PC model compared to the benchmark AO model, we can then conclude that the predictability is occurring at parts of the distribution beyond the conditional mean.

We denote the proposed model by AO-$U|X^i$ where the conditional mean of Y^h_{t+h} follows the AO model as given in (2) and the error, U^{ao}_{t+h}, is made dependent on $\tilde{X}^i_t$[1]. Let the conditional density of U^{ao}_{t+h} be

$$h(u|\tilde{X}^i_t) = \frac{d}{du} H(u|\tilde{X}^i_t) \tag{4}$$

where $H(\cdot|\cdot)$ is the conditional CDF of U^{ao}_{t+h}. Then, we define the forecast density of Y^h_{t+h} via

$$f^i_{t+h|t}(Y^h_{t+h}) = h(U^{ao}_{t+h}|\tilde{X}^i_t). \tag{5}$$

[1]Note that the h-step ahead forecast errors U^{ao}_{t+h} will follow a moving average process of order $h-1$, $MA(h-1)$. In fact, this is also true for the errors of AO, AR and PC models. When estimating forecast densities for all models, we focus on their filtered (or pre-whitened) versions after removing the $MA(h-1)$ serial correlation structure.

We now outline a simple approach to estimate $f^i_{t+h|t}(Y^h_{t+h})$. Note from (5) that it is sufficient to estimate $h(u|\tilde{X}_t)$. One possible method of estimation may be to use Hansen (1994) by assuming a parametric distribution for U^{ao}_{t+h}, and then allowing the higher-order parameters (such as skewness) to depend upon $\tilde{X}_t$. For example, Hong *et al.* (2007) use this approach to estimate forecast densities of (high frequency) exchange rates by assuming a generalized skewed-t for the standardized error distribution, and allow the skewness and kurtosis follow an autoregressive process. Ideally, if $h(u|\tilde{X}_t)$ can be represented by a few dimensional parametric distribution, Hansen's approach can be useful to identify which higher-order dynamics (variance, skewness or kurtosis) are affected by the macroeconomic indicator. Although this direction is worth investigating, we instead use a quantile regression approach which is direct and does not require any parametric assumption.

Denote the $\alpha \in (0,1)$ conditional quantile of U^{ao}_{t+h} conditional on $\tilde{X}^i_t = \tilde{x}^i_t$ by $Q_{t+h}(\alpha|\tilde{x}^i_t)$[2]. We estimate $Q_{t+h}(\alpha|\tilde{x}^i_t)$ using quantile regression model (Koenker and Bassett, 1978),

$$Q_{t+h}(\alpha|\tilde{x}^i_t) = \delta_{0,\alpha} + \sum_{k=1}^{q} \delta_{k,\alpha} x^i_{t-k+1}. \tag{6}$$

Although the local effect of x^i_{t-k+1} on the α-quantile is assumed to be linear, the model is very flexible because each slope coefficient $\delta_{k,\alpha}$ is allowed to differ across quantiles. This is a useful property since it provides guidance as to where in the distribution of Y^h_{t+h} the indicator X^i_t has a significant effect. Of course, the effect of a macro variable X^i_t may well be non-linear. Possible non-linearity can be easily entertained by extending (6) to additive models; see for example, de Gooijer and Zerom (2003), among others. We think that linear quantiles are already flexible enough to capture higher order features of the forecast errors under our set-up. We estimate $h(u|\tilde{X}^i_t)$ using (4) where

$$\hat{H}(u|\tilde{X}^i_t) = \int_0^1 \mathbf{1}\Big(\hat{Q}_{t+h}(\alpha|\tilde{X}^i_t) \leq u\Big) d\alpha \tag{7}$$

with $\mathbf{1}(A)$ denoting an indicator function of set A. An advantage of (7) is that even when the conditional quantile $Q_{t+h}(\alpha|\tilde{X}^i_t)$ may not be monotonic in α, the conditional distribution $\hat{H}(u|\tilde{X}^i_t)$ stays monotonic in u, see Chernozhukov *et al.* (2006) for more details.

[2] One could also estimate the quantiles of inflation directly where various quantiles of Y^h_{t+h} (not U^{ao}_{t+h}) are modeled as linear functions of lagged inflation and activity indicators. Our choice in this paper is guided by the aim of incorporating the existing empirical evidence that favors the naïve model in a more general model with dependence occurring also at higher moments. Consequently, any forecast improvement we report can be attributed to effects observed (or not observed) beyond the conditional mean.

3 Measuring relative predictability

In this paper our focus is on forecasting the distribution of future inflation. Clearly, this raises the question of how to measure the relative accuracy of a particular forecast density as compared to a certain benchmark. Much of the approaches to evaluate forecast densities have mainly focused on their absolute accuracy by developing tests that examine their dynamical and distributional misspecification, see for example Diebold *et al.* (1998) and Hong *et al.* (2007). However, it is very likely that empirical forecasting models are, to some extent, misspecified. In this sense, an "absolute" evaluation measure of one or more forecast densities would not be that informative. On the other hand, we might be willing to accept a possibly misspecified model if it provides a more accurate forecast density relative to another model. This is the approach we take to evaluate the (relative) accuracy of a particular forecast density.

Let's assume there are two forecasting methods used to estimate the density forecast of the h-month ahead inflation, Y^h_{t+h}, where one of them is the benchmark model. As benchmark model, we consider two univariate models, i.e., the AR model (see Equation 3) and the AO model (see Equation 2). We denote the benchmark forecast density by $f^0_{t+h|t}(Y^h_{t+h})$[3] where $0 \in \{\text{AO, AR}\}$. The two benchmark (univariate) forecast densities are separately compared against alternative (multivariate) models that incorporate the effect of macroeconomic indicators. For the latter, we consider density forecasts from the PC model (see Equation 1) where we denote its density by $f^{\text{PC}}_{t+h|t}(Y^h_{t+h})$[4] and AO-$U|X^i$ (denoted by $f^i_{t+h|t}(Y^h_{t+h})$). For the purpose of this Section, let's denote these alternative models by $f^1_{t+h|t}(Y^h_{t+h})$ where $1 \in \{\text{PC}, i\}$.

We adopt a rolling window approach when generating all out-of-sample density forecasts from the various approaches outlined above. Let T be the total number of available observations and t_0 be the first forecast base. This means that there are t_0 observations up to and including the t_0-th observation. By rolling it is meant that the forecast base t extends as far as $T - h$ where h is the forecast horizon. Hence, we have $t = t_0, t_0 + 1, \ldots, T - h$. The goal is to compare the relative accuracy of the two forecast densities even if both models may be misspecified. In other words, which forecast density provide better predictability. We use an intuitively simple metric introduced by Giacomini and White (2006) and Amisano and Giacomini (2007) although other similar suggestions can also be used, see Mitchell and Hall (2005) and Bao *et al.* (2007). This

[3] To ensure consistency, we estimate these densities in a similar fashion as that of $f^i_{t+h|t}(Y^h_{t+h})$ except that the quantile regression in (6) is based on regressing the filtered errors U^{ao}_{t+h} or U^{ar}_{t+h} on only a constant (hence their respective error distribution is assumed to be independent of economic indicators).

[4] Estimated in the same way as that explained in footnote (3) except that the filtered error U^{pc}_{t+h} is used.

metric is based on the average logarithm score of two competing forecast densities defined as follows

$$\frac{1}{N}\sum_{t=t_0}^{T-h}\log f^0_{t+h|t}(Y^h_{t+h}) \quad \text{and} \quad \frac{1}{N}\sum_{t=t_0}^{T-h}\log f^1_{t+h|t}(Y^h_{t+h})$$

where $N = T - h - t_0 + 1$. Note that the forecast densities are evaluated at the realized h-step true value, Y^h_{t+h}. Then, the forecast density (f^0 or f^1) with higher average logarithm score is said to be relatively more accurate. Based on this idea, Amisano and Giacomini (2007) introduce a test procedure to evaluate the null hypothesis of equal density forecast accuracy. Let

$$WLR_t = \log f^0_{t+h|t}(Y^h_{t+h}) - \log f^1_{t+h|t}(Y^h_{t+h}), \quad t = t_0, t_0+1, \cdots, T-h. \tag{8}$$

Note that WLR_t can also be weighted if one is interested to focus only on a certain aspect of the distribution (such as the center or the tails). In this paper, we compare forecast densities in the complete range of variation of the variable and hence weighting is not applied. To test the null hypothesis

$$H_0: \quad f^1(\cdot) \quad \text{and} \quad f^0(\cdot) \quad \text{are equally accurate on average,} \tag{9}$$

we use the AG (Amisano and Giacomini) test statistics, which is defined as

$$\text{AG} = \hat{\sigma}_N^{-1} N^{-\frac{1}{2}}\ \overline{WLR}_N \quad \sim \quad N(0,1)$$

where $\overline{WLR}_N$ is the sample average of WLR_t (over N observations) and $\hat{\sigma}^2_N$ is a heteroskedasticity and autocorrelation consistent (HAC) estimator of the asymptotic variance of $N^{-\frac{1}{2}}\overline{WLR}_N$. Rejections that occur for AG < 0 indicate that $f^1(\cdot)$ is a significantly more accurate density forecast relatively to $f^0(\cdot)$, and vice versa for AG > 0.

3.1 How to interpret the AG test

In order to put the contribution of this paper in the context of the existing empirical evidence on U.S. inflation predictability, we provide a brief discussion on interpreting the AG test. After 1984, the overwhelming majority of empirical studies on forecasting inflation suggests that indicators of economic activity do not carry much relevant information about the conditional mean of the inflation process. Using the relative Root Mean Square Prediction Error for the PC model (Rel-

RMSPE-PC) as a measure of accuracy, this means that

$$\text{Rel-RMSPE-PC} = \frac{\text{RMSPE of PC model}}{\text{RMSPE of AO model}} \geq 1. \tag{10}$$

For these two models (PC and AO), (10) also implies that the AG-test will not lead to rejection of the null hypothesis (see Equation 9) because the only part of the conditional distribution that is allowed to depend on $\tilde{X}_t$ (for the PC model) is the conditional mean. So both the AG-test and Rel-RMSPE-PC will lead to the same conclusion for the PC model (relative to the AO model). On the other hand, the relative RMSPE of the AO-$U|X^i$ model is equal to 1 by construction. Thus, if the forecast density of AO-$U|X^i$ (for some economic indicator) is found to be more accurate than the AO model as reflected in the AG-test (for post 1984 period), this will suggest that $\tilde{X}_t$ carries relevant information for moments beyond the conditional mean.

To summarize, a particular macroeconomic indicator is said to have an effect beyond the conditional mean of the h-step future inflation when the following occur: (a) The Rel-RMSPE-PC is ≥ 1 and (b) The AG-test shows that the forecast density of AO-$U|X^i$ is more accurate than the AO model.

4 U.S. inflation forecasts

We use four measures of the monthly price index (P_t): Consumer Price Index for all items (CPI), CPI excluding food and energy (core-CPI), Personal Consumption Expenditure deflator (PCE), and the PCE excluding food and energy (core-PCE). We follow the recent inflation forecasting literature (see Stock and Watson, 1999, and Ang *et al.*, 2006) and include six of the indicators of economic activity that are often considered as predictors of inflation, i.e., the civilian unemployment rate (UNEM), the index of industrial production (IP), real personal consumption expenditure (INC), employees on non-farm payrolls (WORK), housing starts (HS), and the term spread (SPREAD) defined as the yield on the 5-year Treasury bond minus the 3-month Treasury bill. Thus, we have

$$X_t^i \quad \text{where} \quad i \in \{\text{UNEM, IP, WORK, HS, INC, SPREAD}\}.$$

For both the PC model and the proposed AO-$U|X^i$ model, we find that including lags of X_t^i does not provide forecast improvements over the no lag case. So all results in this Section are based on conditioning on current values of economic indicators, i.e. X_t^i. We implement the AO-$U|X^i$ model by estimating the conditional quantiles for levels α between 0.05 and 0.95 at 0.01 intervals.

All the data (on P_t and X_t^i) were gathered from the Federal Reserve Bank of Saint Louis database FRED and the sample period spans from January 1959 until December 2007[5]. Some of the leading indicators (i.e., IP, INC, and WORK) are not stationary. We thus consider these variables in gap form where the long-run trend is modeled using a Hodrick and Prescott (1997) filter (HP) with parameter equal to 14400 (typically used for monthly data)[6]. The trend is estimated only on information available at the time the forecast is made.

4.1 AG test and Rel-RMSPE results

In the estimation of all forecast densities we use a rolling window scheme as described in Section (3). Our first forecast is January 1985 and the models are estimated on the window 1959:1 to 1984:12 minus the forecasting horizon h (equal to 6 and 12 months[7]). The next forecast is for February 1985 and so on. The size of the rolling window (i.e. 300 monthly observations) is kept constant by dropping one observation at the beginning of the sample. For the AR and PC models, we select the lag order for ΔY_t using the Akaike Information Criterion (AIC), recursively for each rolling window sample. It should be noted that the forecast evaluation period is chosen intentionally to be post 1984, when current research shows that macroeconomic indicators do not add much to the predictability of mean inflation. We also divide our post 1984 evaluation period into two equal sub-periods (1985:1 to 1996:6 and from 1996:7 to 2007:12, each sub-period consisting of 138 months) to examine if there was any significant change in predictive power of macroeconomic variables.

We report results for CPI and PCE in Table (1) and Table (2), that correspond to $h = 12$ and $h = 6$, respectively. Results for core-CPI and core-PCE are shown in Table (3) and Table (4).

CPI and PCE

For both $h = 12$ and $h = 6$, the Rel-RMSPE-PC is greater than 1 (except when using HS and SPREAD for which the ratio is barely less than 1), indicating that the naïve model outperforms the PC model. This simply confirms the earlier results in the literature on the inability of activity indicators to predict mean inflation in post 1984 period. We also report Rel-RMSPE-AR which is consistently larger than 1. Examining the AG test results, we observe (for both $h = 12$ and

[5]The macroeconomic series consists of revised data available at the January 2008 vintage due to the lack of a comprehensive real-time dataset at the monthly frequency.

[6]We also considered a quadratic trend as in Ang *et al.* (2006) but the results are very similar to the HP filter. To conserve space we decided to report only the results of the HP filter.

[7]We also used a one quarter horizon. However, the results were largely similar to the semi-annual and annual horizon and decided not to report them in this paper.

$h = 6$) that none of the macroeconomic variables have any significant predictive power for CPI. Therefore, the forecast densities of the AO-$U|X^i$ model (for all X^i) are not more accurate than the AO model for inflation measured by CPI. In contrast, HS and SPREAD seem to provide some significant density predictability of inflation when PCE is used although this effect is limited to the first sub-period 1985:1-1996:6. In the second sub-period 1996:7-2007:12, there is also evidence of better density predictability of PCE when using IP GAP, INC GAP and SPREAD (see that all the AG statistics are negative and relatively large) although these better performances are not statistically significant.

Core-CPI and Core-PCE

For both $h = 12$ and $h = 6$, the Rel-RMSPE-PC is greater than 1 (except when using HS and SPREAD for which the ratio is less than 1 especially in the second sub-period 1996-2007). With few exceptions, this result again confirms the earlier findings in the literature on the inability of activity indicators to predict mean inflation for the post 1984 period. When examining the AG tests, the results are very different from those found for CPI and PCE. When inflation is measured by core-CPI and core-PCE, there is significant evidence of density predictability in both sub-periods and horizons. At the annual horizon ($h = 12$), a large number of activity indicators are useful in providing more accurate forecasts of the distribution of inflation for core-PCE (in particular in the second sub-period). In predicting core-CPI at $h = 12$, UNEM, HS, and SPREAD are found to be useful in the first sub-period and UNEM provides better predictability also in the second sub-period. The AG test results are even stronger for $h = 6$. For example, considering core-CPI, all economic indicators are found to be significant in the first sub-period. The usefulness of half of the indicators also extends to the second sub-period. For core-PCE, predictive relevance of the indicator variables appears to be concentrated in the second sub-period.

4.2 Fluctuation test results

A conclusion that emerges from the foregoing out-of-sample predictability evidence (for the post 1984 period) based on the AG test results is that some economic indicators provide improvements in accuracy of the density forecasts for core-CPI and core-PCE, although they seem to be weakly informative when forecasting CPI and PCE. In this Section, an attempt is made to shade some light on the reasons behind the marked differences in forecast performances observed across various measures of inflation. We accomplish this by assessing relative forecast performances using the fluctuation test of Giacomini and Rossi (2010)[8].

The relative forecast performance results reported in Table (1)-(4) are based on the AG statistic, which consists of averaging (aggregating) performances over a sub-period of 11.5 years, i.e. 1985:1-1996:6 and 1996:7-2007:12. So, the AG test will surely overlook valuable relative local performances that occur at certain times. To complement the AG test results, we implement the fluctuation test of Giacomini and Rossi (2010) that evaluates the (relative) performance of competing density forecast models and signals whether one of the densities is significantly (given appropriate critical values) more accurate than the other density at any point in time. Unlike the AG test, which evaluates the null hypothesis of density forecast equality on average, the null hypothesis of the fluctuation test is

$$H_0: \quad f_t^1(\cdot) \quad \text{and} \quad f_t^0(\cdot) \quad \text{are equally accurate at each time } t. \tag{11}$$

The fluctuation test is defined as follows (we focus on $h = 12$). We will refer to definitions and notations in Section (3). Based on WLR_t as given in Equation (8), the fluctuation test statistic is defined using two-sided rolling window of m observations,

$$FL_t = \hat{\sigma}_m^{-1} \; m^{-1/2} \sum_{j=t-m/2}^{t+m/2-1} WLR_j, \qquad t = t_0 + m/2, \cdots, T - h - m/2 + 1$$

where $\hat{\sigma}_m^2$ is a heteroskedasticity and autocorrelation consistent (HAC) estimator of the asymptotic variance. The asymptotic distribution of FL_t under the null hypothesis is non standard. Critical values for various $\mu = m/N$ and two significance levels are provided in Giacomini and Rossi (2010). In our case, we have $N = 276$ and $m = 120$ resulting in $\mu \approx 0.43$. The null hypothesis in (11) is rejected at the 5% (10%) level against a two-sided alternative when $\max_t |FL_t| > 2.890$ (2.626). Note that FL_t over time also contains valuable information. If the FL_t crosses the lower bound at time t then we conclude that $f_t^1(\cdot)$ is more accurate, whereas if FL_t crosses the upper bound we conclude that $f_t^0(\cdot)$ is more accurate. Here we focus on comparing the forecast density $f_t^1(\cdot)$ of the AO-$U|X^i$ model against the forecast density $f_t^0(\cdot)$ of the benchmark AO model and discuss the fluctuation test results for all four inflation measures. The fluctuation test results for CPI and PCE are given in Figure (1) whereas those for core-CPI and core-PCE are shown in Figure (2).

[8]See also Rossi and Sekhposyan (2010) for a recent application of this approach to compare (conditional mean) forecasting performance of different models for US output and inflation.

Focusing on Figure (1a) which corresponds to CPI, we can observe that there is evidence of predictability based on most indicators (the FL_t statistic takes a negative value) in the period 1997-2007, although none is statistically significant at conventional levels. In addition, the fluctuation statistic for INC GAP and HS is negative in most of the period considered indicating that the AO-$U|X^i$ model forecast density is more accurate (although not significant) than the AO model. For PCE (see Figure 1b), the pattern of the FL_t statistic is similar to CPI with better predictability (compared to the AO model) after 1997 for most economic indicators. This result explains the modestly large negative AG statistics values shown in Table (1) for all indicators. The best performance is achieved by HS for which FL_t is negative and significant at 10% (based on the two-sided critical values) in the periods 1990-1992 and between mid-1995 and 1997.

Moving to the core inflation measures, Figure (2a) shows the time path of FL_t of core-CPI. The FL_t pattern for core-CPI shows a much better performance than that for CPI especially before 1997 where almost all the FL_t are consistently negative. For UNEM, HS, and SPREAD, these FL_t are also significant for some periods which explain the 1985:1-1996:6 AG test results of core-CPI reported in Table (3). In the Table it is shown that UNEM, HS and SPREAD provide significant predictability for core-CPI. Comparing Figure (2a) with Figure (1a), the most striking difference is observed for UNEM and HS where these economic indicators become relevant in forecasting core-CPI but are irrelevant to forecast CPI. Notice from Figure (2a) that the FL_t for both UNEM and HS are uniformly negative for the whole post 1984 period.

Figure (2b) shows the time path of FL_t of core-PCE. This measure of inflation is the one with the most predictability. Interestingly, all the economic indicators provide predictability (see the negative FL_t for all of them) during the recent period 1996-2007 and both UNEM and WORK GAP provide significant predictability throughout this period. This result is consistent with Table (3) where core-PCE is found to be significantly predictable by all six economic indicators during the 1996-2007 period. A further look at Figure (2b) indicates that UNEM, HS and SPREAD have negative FL_t through out the post 1984 period and the FL_t for UNEM and HS are also significant for large part of 1985-2007.

To summarize, there is ample evidence of predictability of inflation for all versions of inflation although the evidence is stronger for the core versions of inflation, i.e. core-CPI and core-PCE. In particular, the core-PCE measure of inflation shows the most significant evidence of predictability notably when using UNEM and HS as predictors throughout the 1985-2007 period. The fluctuation test is a useful tool in uncovering relevant predictability information that is overlooked in average measures such as the AG test. Relying only on the AG test will give the impression that all the considered economic indicators are uninformative in predicting CPI and PCE versions of inflation. But a much more detailed predictability picture provided by the fluctuation test shows that many of the indicators are indeed relevant although better performances seem to be overwhelmed by periods of poor performance which on aggregate result in poor AG values.

4.3 Distribution forecasts of inflation

In this paper we ask the question: are macroeconomic variables useful in forecasting the distribution (beyond the mean) of U.S. inflation in the post 1984 period? Using AG and fluctuation tests, we find that some of the economic indicators we consider in this paper provide significant predictability of the distribution of inflation (especially for core-PCE) for large part of the post 1984 period.

To offer insights as to why the distribution of inflation is found predictable (see the AG and fluctuation test results), we examine selected quantiles of the h-step ahead out-of-sample forecast distribution of inflation (core-PCE). We consider 5%, 50% and 95% quantile levels that represent the lower tail, the middle and upper tail of the forecast distribution, respectively. The 50% quantile is the median forecast which will coincide with the usual mean forecast when the forecast distribution is symmetric. Because of the compelling evidence in favor of HS and UNEM in forecasting the distribution of core-PCE, we focus on these two economic indicators. To conserve space, we restrict the analysis to the case of $h = 12$.

Housing Starts (HS)

In the top portion of Figure (3), we show the quantile forecasts of core-PCE for the 1985:1-1996:6 period. We also include the corresponding quantile forecasts of the benchmark AO model (shown using broken lines). In the lower panel of the Figure we display the time series plot of HS which is shifted forward by 12 so that it becomes aligned with the forecasting date. For example, the HS value in January 1985 actually refers to that of January 1984 which represents the value of HS used to produce the distribution forecasts for the target date (January 1985). For the 1985:1-1996:1 period, the AG test (see Table 3) shows that HS provides significant predictability in the distribution of core-PCE. This result is a reflection of the fluctuation test results reported in Figure (2b) where we observe that HS provides significant predictability in that period. However, notice that the Rel-RMSPE-PC (see Table 3) is 1.012 suggesting HS does not help predict the conditional mean of core-PCE. Taken together, the above results imply that the predictability by

HS occurs in parts of the conditional distribution of core-PCE other than the mean. Now looking at Figure (3), notice that, for most of this period, the median forecasts based on HS closely track those of the AO model which may explain the inability of HS to predict mean inflation (assuming mean and median forecasts are comparable). The same pattern is also observed for the lower tail quantile forecasts (the 5% level). In contrast, the upper tail (95%) quantile forecasts based on HS are much lower than the corresponding quantile of the AO model. These lower and upper tail quantile forecast patterns imply that the higher accuracy of distribution forecasts based on HS during 1985:1-1996:6 is the result of better predictability occurring at the higher tail quantile. Notice from the time plot of HS (see the lower portion of Figure 3) that the slow and steady decrease in housing starts (from mid-1980's to beginning of 1990) exerts the strongest downward pressure on high levels of inflation.

Unlike 1985:1-1996:6 period, the Rel-RMSPE-PC (see Table 3) for 1996-2007 period is 0.908 implying that HS does help predict the conditional mean of core-PCE. Moreover, for the latter period both the AG (see Table 3) and fluctuation tests (see Figure 2b) indicate that HS provide significant predictability for the distribution of core-PCE. From these observations, one may raise the question: can we attribute the distribution predictability provided by HS (for 1996-2007) solely to predictability of the mean? In the top portion of Figure (4), we show the quantile forecasts of core-PCE for the 1996-2007 period together with the corresponding quantile forecasts of the benchmark AO model (shown using broken lines). Looking at Figure (3), notice that the median forecasts based on HS are very different from those of the AO model. It can also be observed that both tail quantile forecasts (at 5% and 95%) are different from the AO model although the difference is larger at the lower tail quantile. So, the better distribution forecasts observed for HS during 1996 - 2007 can not be solely explained by better mean forecasts as there appears to be better predictability occurring at the lower tail quantile. Notice from the time plot of HS (shown in the lower portion of Figure 3) that the slow and steady increase in housing starts (from 2001 to 2007) exerts an upward pressure on the whole distribution of inflation although this pressure seems to be more pronounced at the lower tail quantile.

Unemployment rate (UNEM)

From Table (3), for both periods 1985:1-1996:6 and 1996:7-2007:12 the Rel-RMSPE-PC is higher than 1 (1.043 and 1.129, respectively) which suggests that UNEM does not help predict the conditional mean of core-PCE. In contrast, the AG and fluctuation tests (see Figure 2b) show that UNEM provides significant predictability in the distribution of core-PCE in both periods.

So, we may infer that the distribution predictability provided by UNEM is attributed to effects that occur at parts of the distribution other than the mean. To explore this result further, we will focus on the quantile forecasts for 1996:7-2007:12 period. Results for 1985:1-1996:6 are similar.

In the top portion of Figure (5), we show the quantile forecasts of core-PCE for the 1996:7-2007:12 period when using the UNEM predictor. Looking at Figure (5), notice that UNEM-based quantile forecasts at all levels (5%, 50% and 95%) are different from the corresponding quantile forecasts of the benchmark AO model. It appears that the conditional distribution of core-PCE that is based on UNEM is shifted upward compared to the AO model (shown using broken lines). Note that the discrepancy between the two quantile forecasts is a lot larger at the lower tail (5%) quantile level where the quantile forecasts based on UNEM are much higher than the corresponding quantile forecasts of the AO model. This result may be attributed to the upward pressure on inflation derived from the persistent decrease in the unemployment rate in the late 1990's as observed from the time series plot of UNEM (see the lower portion of Figure 5). In the late 1990's and early 2000 unemployment was at historically low levels approaching 4%. Consistent with the existence of a trade-off between unemployment and inflation rates, the low unemployment rate shifted the whole forecast distribution of inflation to higher levels (compared to the AO model distribution forecasts) and the relative shift is much larger at the lower tail quantile of the distribution.

4.3.1 Forecasting the probability of deflation

In the beginning of 1998 a debate started on the possibility of the U.S. economy entering a period of deflation that was also discussed in a speech given by the Federal Reserve Board Chairman on January 3rd, 1998 (see Greenspan, 1998). Using this historical fact as a motivation, we now present a complementary (to AG and Fluctuation tests) approach to evaluate forecast distributions of inflation by focusing on the accuracy of the methods in predicting core-PCE deflation (negative inflation). We compare the probability of deflation of the benchmark AO model to that of AO-$U|X^i$ where the quantiles are conditional on $i = UNEM$. We consider the case of $h = 12$ (1 year-ahead forecast) which is often of particular interest. In Figure (6) we show the probability forecasts of deflation with and without UNEM for the period 1996:7-2007:12. From the Figure it can be observed that a forecaster using the AO model would have predicted the probability of deflation to be larger than 5% for several months over the years with the highest probability reaching 18.2% for the June 1999 target date. In contrast, for the same June 1999, AO-$U|X^{UNEM}$ would have predicted only a 1.18% chance of deflation. Overall, the use of the unemployment rate leads to a likelihood of deflation close to zero throughout the 1996:7-2007:12 period, and thus appears to provide a more appropriate prediction of the event.

5 Conclusion

Forecasting the behavior of inflation plays a central role in economic policy-making due to the inherently forward-looking nature of economic decisions. Typically, inflation forecasting focuses on modeling the conditional mean or the most likely outcome. While relevant, relying only on the dynamics of the conditional mean leaves out other interesting aspects of the inflation process, such as the dynamics occurring at higher moments of the inflation distribution. For example, the central bank may be interested in evaluating "unattractive" outcomes for the economy such as deflation or high inflation, and those models that rely only on the conditional mean will not offer the tools to do such evaluations. Accurate characterization of the complete distribution of future inflation, beyond the conditional mean, is needed.

This paper examines whether indicators of economic activity carry relevant information about the dynamics of higher moments of inflation, and hence help improve the accuracy of density forecasts. Our findings indicate that, in particular for the core inflation measures, conditioning the dynamics of the inflation distribution on the leading indicators provides more accurate forecasts relative to the random walk model. This is due to the relevance of the activity indicators in forecasting quantile effects that take place far away from the center of the core inflation distribution. We also investigate the possibility of episodic predictability, in the sense that economic indicators might provide more accurate density forecasts during limited periods of the forecasting sample. The results indicate that some variables (in particular housing starts and the unemployment rate) have this characteristic, in particular when forecasting PCE and core-CPI, while they provide consistent evidence of predictability for core-PCE throughout the period 1985-2007.

Overall, our results indicate that economic variables are more useful indicators of the dynamics of the tails of the inflation distribution, rather than its center. This finding can be of particular interest for policy makers when evaluating the likelihood of certain events, such as whether inflation will be above or below a certain level in the future.

References

Amisano, G. and Giacomini, R. (2007). Comparing density forecasts via Weighted Likelihood Ratio Tests. *Journal of Business & Economic Statistics*, **25,** 177–190.

Ang, A., Bekaert, G. and Wei, M. (2006). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, **54,** 1163–1212.

Atkenson, A. and Ohanian, L. (2001). Are Phillips Curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, **25,** 2–11.

Bao, Y., Lee, T.-H. and Saltoglu, B. (2007). Comparing density forecast models. *Journal of Forecasting*, **26,** 203–225.

Chernozhukov, V., Fernandez-Val, I. and Galichon, A. (2006). Quantile and probability curves without crossing. *working paper.*

Clark, T.E. and McCracken, M.W. (2006). The predictive content of the output gap for inflation: resolving in-sample and out-of-sample evidence. *Journal of Money, Credit, and Banking*, **38,** 1127–1148.

Cogley, T., Morozov, S. and Sargent, T.J. (2005). Bayesian fan charts for U.K. inflation: forecasting and sources of uncertainty in an evolving monetary system. *Journal of Economic Dynamics and Control*, **29,** 1893–1925.

Corradi, V. and Swanson, N.R. (2006). Predictive density and conditional confidence interval accuracy tests. *Journal of Econometrics*, **135,** 187–228.

de Gooijer, J.G. and Zerom, D. (2003). On additive conditional quantiles with high-dimensional covariates. *Journal of the American Statistical Association*, **98,** 135–146.

Diebold, F.X., Gunther, T.A. and Tay, A.S. (1998). Evaluating density forecasts with application to financial risk management. *International Economic Review*, **39,** 863–883.

Fisher, J.D.M., Liu, C.T. and Zhou, R. (2002). When can we forecast inflation? *Federal Reserve Bank of Chicago Economic Perspectives* 30–42.

Giacomini, R. and Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics,* forthcoming.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, **74,** 1545–1578.

Greenspan, A. (1998). Problems of price measurement. *Federal Reserve Board, Testimony and Speeches (January 3rd).*

Greenspan, A. (2004). Risk and uncertainty in monetary policy. *American Economic Review*, **94,** 33–40.

Hodrick, R.J. and Prescott, E.C. (1997). Postwar U.S. business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, **29,** 1–16.

Hong, Y., Li, H. and Zhao, F. (2007). Can the random walk model be beaten in out-of-sample density forecasts? evidence from intraday foreign exchange rates. *Journal of Econometrics*, **141,** 736–776.

Inoue, A. and Kilian, L. (2004). In-sample or out-of-sample tests of predictability? Which one should we use? *Econometric Reviews*, **23,** 371–402.

Kilian, L. and Manganelli, S. (2008). The central banker as a risk manager: estimating the Federal Reserve's preferences under Greenspan. *Journal of Money, Credit, and Banking*, **40,** 1103–1129.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, **46,** 33–50.

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 Bulleting of Economics and Statistics*, **67,** 995–1003.

Robertson, J.C., Tallman, E.W. and Whiteman, C.H. (2005). Forecasting using relative entropy. *Journal of Money, Credit, and Banking*, **37,** 383–401.

Rossi, B. and Sekhposyan, T. (2010). Has models' forecasting performance changed over time, and when? *International Journal of Forecasting,* forthcoming.

Stock, J.H. and Watson, M.W. (1999). Forecasting inflation. *Journal of Monetary Economics*, **44,** 293–335.

Stock, J.H. and Watson, M.W. (2007). Why has U.S. inflation become harder to forecast. *Journal of Money, Credit, and Banking*, **39,** 3–33.

Stock, J.H. and Watson, M.W. (2008). Phillips curve inflation forecasts. *working paper*.

Table 1: $\mathbf{h = 12}$

Variable i	**Method**	**CPI**						**PCE**					
		1985:1-1996:6			1996:7-2007:12			1985:1-1996:6			1996:7-2007:12		
		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test	
			AR	AO		AR	AO		AR	AO		AR	AO
	AR	1.131		1.414	1.029		0.504	1.202		3.826	1.05		0.484
UNEM	PC	0.992	-0.107	1.651	1.050	0.786	0.805	1.010	-0.549	3.986	1.033	0.513	0.573
	AO-$U\|X^i$		-0.898	0.407		-0.156	0.261		**-2.112**	0.162		-0.875	-0.512
IP GAP	PC	1.061	0.898	1.637	0.991	-0.294	0.365	1.061	1.298	3.569	0.952	-0.020	0.520
	AO-$U\|X^i$		-0.372	1.253		-0.716	-0.677		**-2.626**	0.531		-1.380	-1.314
INC GAP	PC	1.085	1.274	1.913	0.951	**-1.99**	-0.276	1.101	0.870	2.803	0.989	-0.443	0.368
	AO-$U\|X^i$		-0.161	0.800		-1.000	-1.251		**-2.489**	0.127		-1.410	-1.403
WORK GAP	PC	1.022	0.062	1.295	1.049	0.186	0.548	1.039	0.775	3.450	1.044	0.642	0.741
	AO-$U\|X^i$		-0.837	0.630		-0.385	-0.023		**-2.471**	0.559		-1.250	-1.200
HS	PC	1.141	1.444	1.756	0.963	0.046	-0.595	1.065	-0.012	2.352	0.959	-0.080	0.260
	AO-$U\|X^i$		-0.856	0.061		-0.747	0.530		**-10.81**	**-2.875**		-1.493	-1.289
SPREAD	PC	1.072	1.528	1.726	0.988	-0.447	0.453	1.032	1.956	3.981	0.993	-0.431	0.404
	AO-$U\|X^i$		0.521	1.614		-1.161	-0.999		**-2.441**	-0.144		-1.597	-1.194

Rel-RMSPE values are calculated relative to the RMSPE of the AO model. For the AO-$U|X^i$ models, we omit the Rel-RMSPE values as they are equal to 1 by construction. AG test indicates the Amisano-Giacomini test and values in bold denote the (one-sided) rejections of the hypothesis of higher accuracy of the alternative model (PC or AO-$U|X^i$) compared to the benchmark model (AR or AO) at 5% significance level.

Table 2: $\mathbf{h = 6}$

Variable i	**Method**	**CPI**						**PCE**					
		1985:1-1996:6			1996:7-2007:12			1985:1-1996:6			1996:7-2007:12		
		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test	
			AR	AO		AR	AO		AR	AO		AR	AO
	AR	1.006		-0.084	0.966		-0.660	1.092		2.727	1.070		1.218
UNEM	PC	1.007	1.145	0.329	1.030	1.436	-0.194	0.997	-0.646	2.575	1.002	0.352	1.259
	AO-$U\|X^i$		0.547	0.969		0.716	0.516		-1.184	0.484		-0.110	0.886
IP GAP	PC	1.084	1.132	0.720	0.990	0.058	-0.639	1.013	-0.246	1.877	1.004	-0.375	1.094
	AO-$U\|X^i$		0.916	1.846		0.699	0.570		**-2.011**	-0.065		-0.142	0.805
INC GAP	PC	1.051	0.669	0.512	0.968	-1.103	-0.994	1.002	-0.654	1.042	0.985	-0.533	0.977
	AO-$U\|X^i$		0.567	0.747		0.657	0.483		**-2.824**	-0.704		-0.506	0.440
WORK GAP	PC	1.079	0.998	0.540	1.033	0.273	-0.477	1.014	-0.414	1.959	1.013	0.289	1.198
	AO-$U\|X^i$		0.855	1.709		0.871	0.962		**-2.143**	-0.128		-0.188	0.868
HS	PC	1.104	1.077	0.939	0.980	-0.359	-0.918	1.008	-0.408	1.215	0.952	-0.745	0.553
	AO-$U\|X^i$		0.104	0.086		0.450	-0.013		**-4.588**	**-1.720**		-1.106	-0.438
SPREAD	PC	0.995	-0.301	-0.140	1.004	0.943	-0.623	0.996	**-1.743**	2.579	0.993	-0.236	1.243
	AO-$U\|X^i$		0.720	1.100		0.544	-0.171		**-2.688**	**-2.032**		-0.217	1.293

Rel-RMSPE values are calculated relative to the RMSPE of the AO model. For the AO-$U|X^i$ models, we omit the Rel-RMSPE values as they are equal to 1 by construction. AG test indicates the Amisano-Giacomini test and values in bold denote the (one-sided) rejections of the hypothesis of higher accuracy of the alternative model (PC or AO-$U|X^i$) compared to the benchmark model (AR or AO) at 5% significance level.

Table 3: $\mathbf{h = 12}$

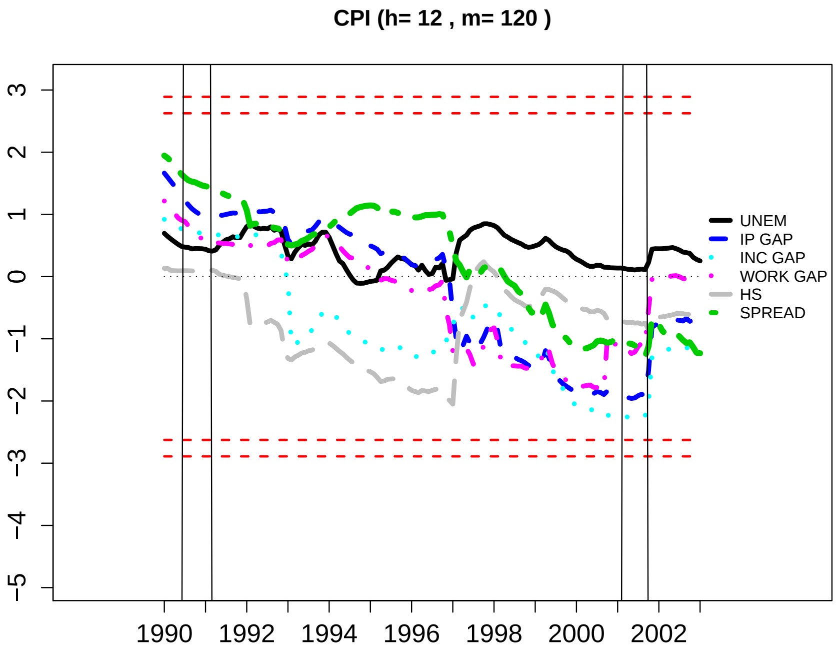
Variable i	**Method**	**core-CPI**						**core-PCE**					
		1985:1-1996:6			1996:7-2007:12			1985:1-1996:6			1996:7-2007:12		
		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test	
			AR	AO		AR	AO		AR	AO		AR	AO
	AR	1.064		9.765	1.175		4.578	1.248		13.47	1.287		2.637
UNEM	PC	0.963	**-6.998**	7.216	1.225	-0.218	4.135	1.043	-0.736	6.663	1.129	0.578	1.917
	AO-$U\|X^i$		**-8.975**	**-3.342**		**-5.223**	**-3.297**		**-14.753**	**-4.646**		**-8.565**	**-4.596**
IP GAP	PC	1.666	-0.696	4.703	1.089	-1.840	3.898	1.168	1.445	8.987	1.024	0.242	2.695
	AO-$U\|X^i$		**-6.959**	0.030		-4.018	-0.852		**-6.276**	-0.817		**-7.810**	**-4.153**
INC GAP	PC	1.449	**-3.031**	5.452	0.987	**-3.943**	3.599	1.145	1.171	15.106	1.000	0.383	2.752
	AO-$U\|X^i$		**-11.456**	-1.104		**-5.207**	-1.580		**-4.814**	-0.823		**-8.195**	**-4.034**
WORK GAP	PC	1.654	-0.586	3.747	1.001	**-5.353**	2.987	1.19	0.92	6.43	1.03	-0.36	2.62
	AO-$U\|X^i$		**-7.576**	0.042		**-4.199**	-1.042		**-6.789**	-0.79		**-8.698**	**-4.916**
HS	PC	1.441	**-3.005**	7.280	0.997	**-2.056**	1.695	1.012	-0.917	7.145	0.908	-1.044	1.102
	AO-$U\|X^i$		**-19.442**	**-2.615**		**-3.964**	-1.213		**-16.918**	**-6.129**		**-7.536**	**-3.714**
SPREAD	PC	0.946	**-6.259**	7.306	1.012	-0.269	4.663	0.992	**-1.824**	7.448	0.962	-1.413	2.329
	AO-$U\|X^i$		**-8.495**	**-3.583**		**-4.374**	-0.291		**-9.055**	**-3.282**		**-7.588**	**-2.934**

Rel-RMSPE values are calculated relative to the RMSPE of the AO model. For the AO-$U|X^i$ models, we omit the Rel-RMSPE values as they are equal to 1 by construction. AG test indicates the Amisano-Giacomini test and values in bold denote the (one-sided) rejections of the hypothesis of higher accuracy of the alternative model (PC or AO-$U|X^i$) compared to the benchmark model (AR or AO) at 5% significance level.

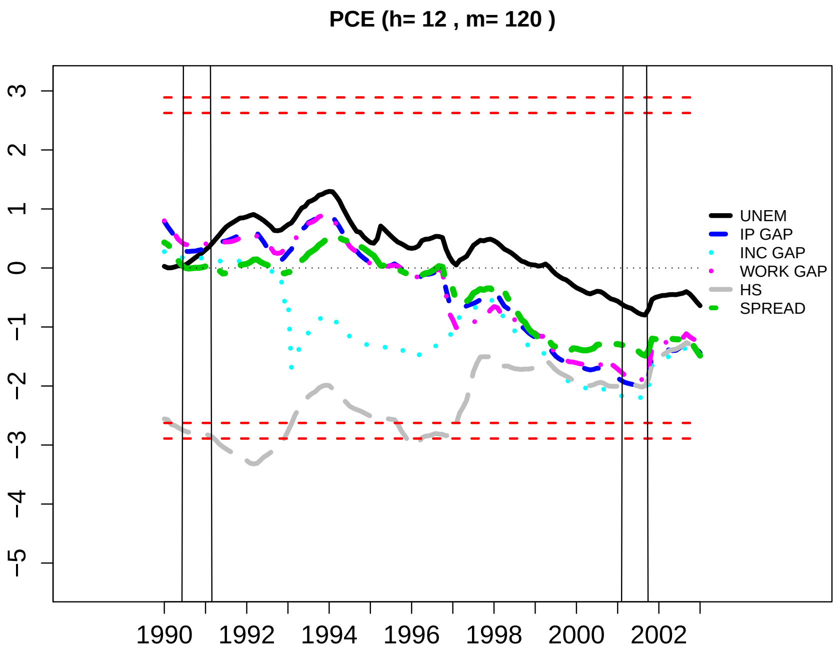
Table 4: $\mathbf{h = 6}$

Variable i	**Method**	**core-CPI**						**core-PCE**					
		1985:1-1996:6			1996:7-2007:12			1985:1-1996:6			1996:7-2007:12		
		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test		Rel-RMSPE	AG test	
			AR	AO		AR	AO		AR	AO		AR	AO
	AR	1.032		11.946	1.054		4.419	1.179		8.613	1.126		1.204
UNEM	PC	0.997	**-2.388**	7.687	1.247	1.814	3.715	1.018	-0.267	5.976	1.055	0.315	0.992
	AO-$U\|X^i$		**-11.435**	**-4.474**		**-4.25**	**-2.507**		**-8.381**	**-3.091**		**-4.265**	**-3.719**
IP GAP	PC	1.517	0.815	5.605	1.128	0.973	3.875	1.048	0.236	7.317	1.015	0.327	1.252
	AO-$U\|X^i$		**-9.799**	**-2.252**		**-4.651**	-1.627		**-4.13**	-0.444		**-4.892**	**-5.574**
INC GAP	PC	1.268	-0.104	7.144	1.016	-1.263	3.476	1.028	0.199	8.417	1.005	0.743	1.389
	AO-$U\|X^i$		**-11.555**	**-1.826**		**-5.529**	**-2.54**		**-5.44**	-1.249		**-4.075**	**-4.347**
WORK GAP	PC	1.598	0.191	3.805	1.059	-1.183	2.896	1.070	0.414	6.565	1.009	-0.724	1.049
	AO-$U\|X^i$		**-10.916**	**-1.866**		**-5.478**	**-2.278**		**-4.019**	-0.387		**-5.143**	**-6.013**
HS	PC	1.272	-0.142	8.995	1.041	-0.815	2.184	0.997	-0.968	7.661	0.934	-1.259	0.187
	AO-$U\|X^i$		**-15.165**	**-3.206**		**-3.682**	-1.271		**-8.606**	**-3.885**		**-3.813**	**-3.339**
SPREAD	PC	1.079	**-1.959**	8.061	1.022	0.209	4.258	1.002	-0.251	6.118	0.984	-0.634	1.042
	AO-$U\|X^i$		**-8.502**	**-4.204**		**-3.957**	0.489		**-9.459**	**-5.426**		**-4.486**	**-4.189**

Rel-RMSPE values are calculated relative to the RMSPE of the AO model. For the AO-$U|X^i$ models, we omit the Rel-RMSPE values as they are equal to 1 by construction. AG test indicates the Amisano-Giacomini test and values in bold denote the (one-sided) rejections of the hypothesis of higher accuracy of the alternative model (PC or AO-$U|X^i$) compared to the benchmark model (AR or AO) at 5% significance level.

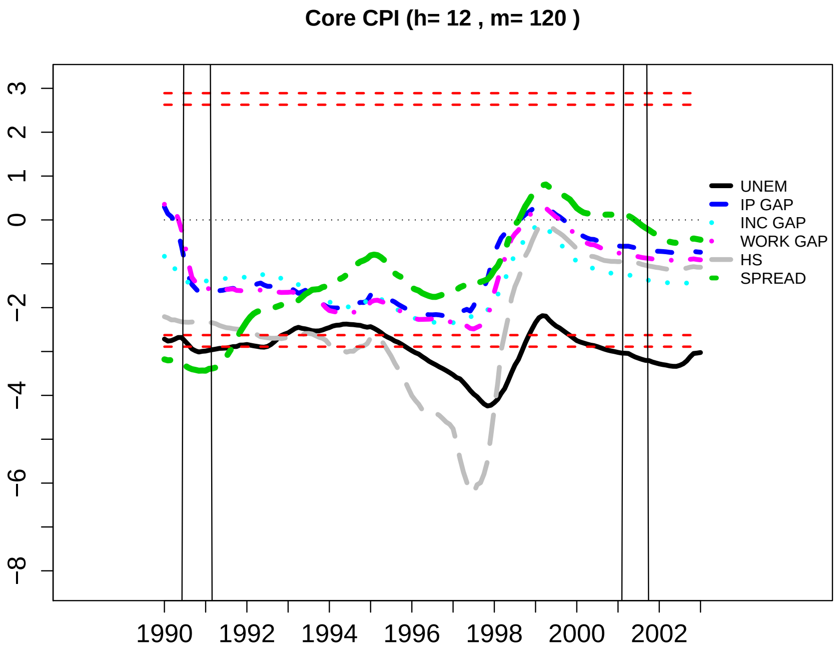


(a)

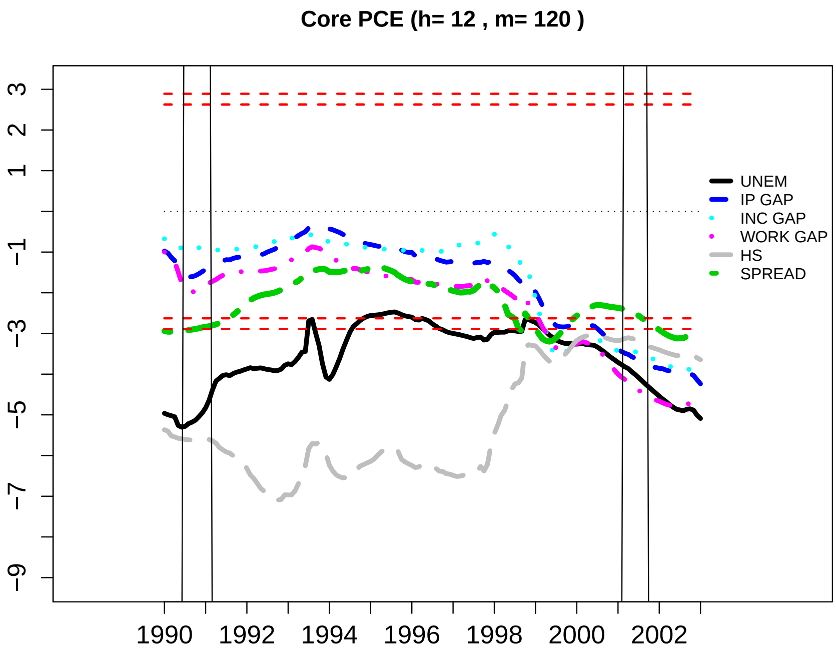


(b)

Figure 1: Fluctuation test for CPI and PCE inflation for AO-$U|X^i$ (i=UNEM, IP GAP, INC GAP, HS, SPREAD) against the benchmark AO model (window size m equal to 120 observations). The dashed horizontal lines represent the 5% and 10% (two-sided) critical values tabulated in Giacomini and Rossi (2010).



(a)



(b)

Figure 2: Fluctuation test for core-CPI and core-PCE inflation for AO-$U|X^i$ (i=UNEM, IP GAP, INC GAP, HS, SPREAD) against the benchmark AO model (window size m equal to 120 observations). The dashed horizontal lines represent the 5% and 10% (two-sided) critical values tabulated in Giacomini and Rossi (2010).

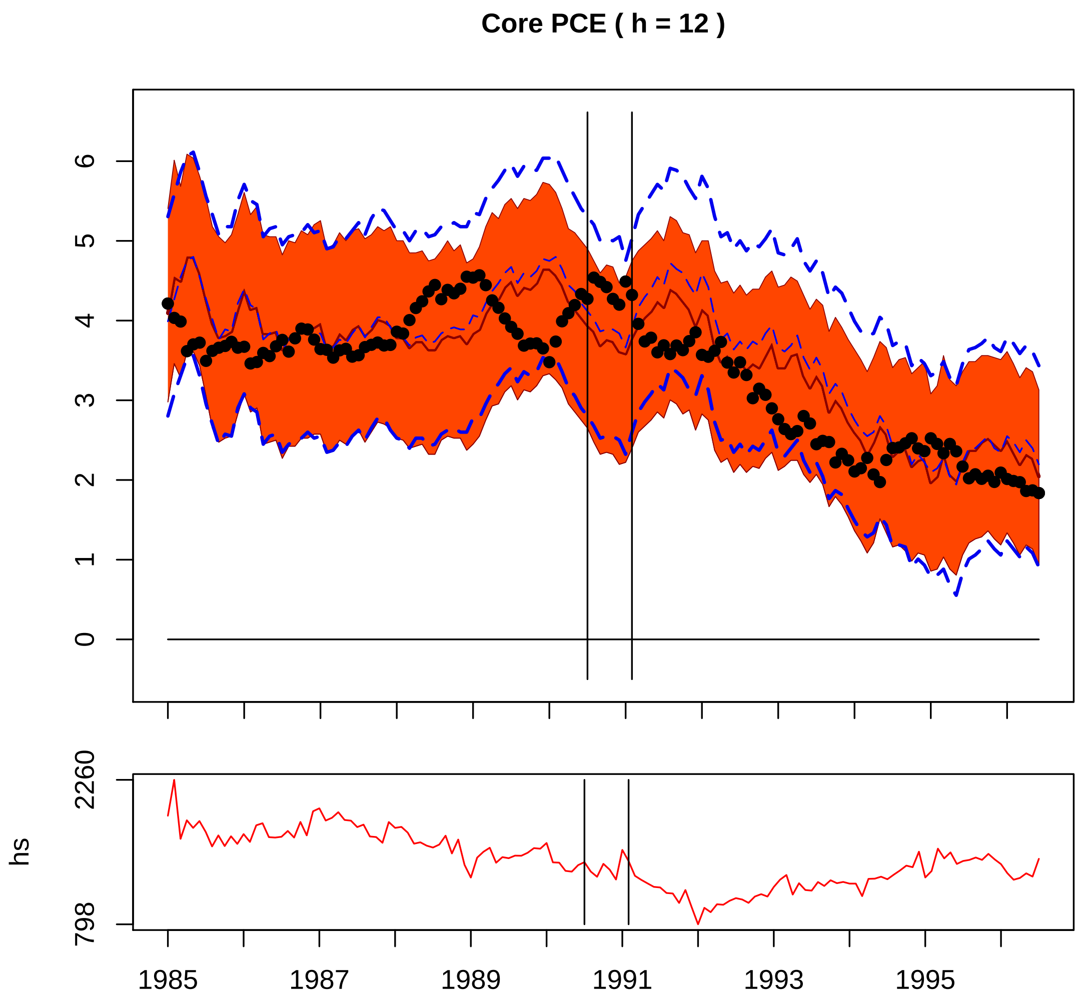


Figure 3: **(Top plot)** The (red) continuous lines denote the quantiles of AO-$U|X^{HS}$ for core-PCE at 5, 50, and 95% levels, the (blue) dashed lines denote the same quantile levels for the AO model, while the circles indicated the realized core-PCE inflation rate. **(Bottom plot)** The housing starts series (shifted forward by $h = 12$ months). The vertical lines indicate the NBER recession dates.

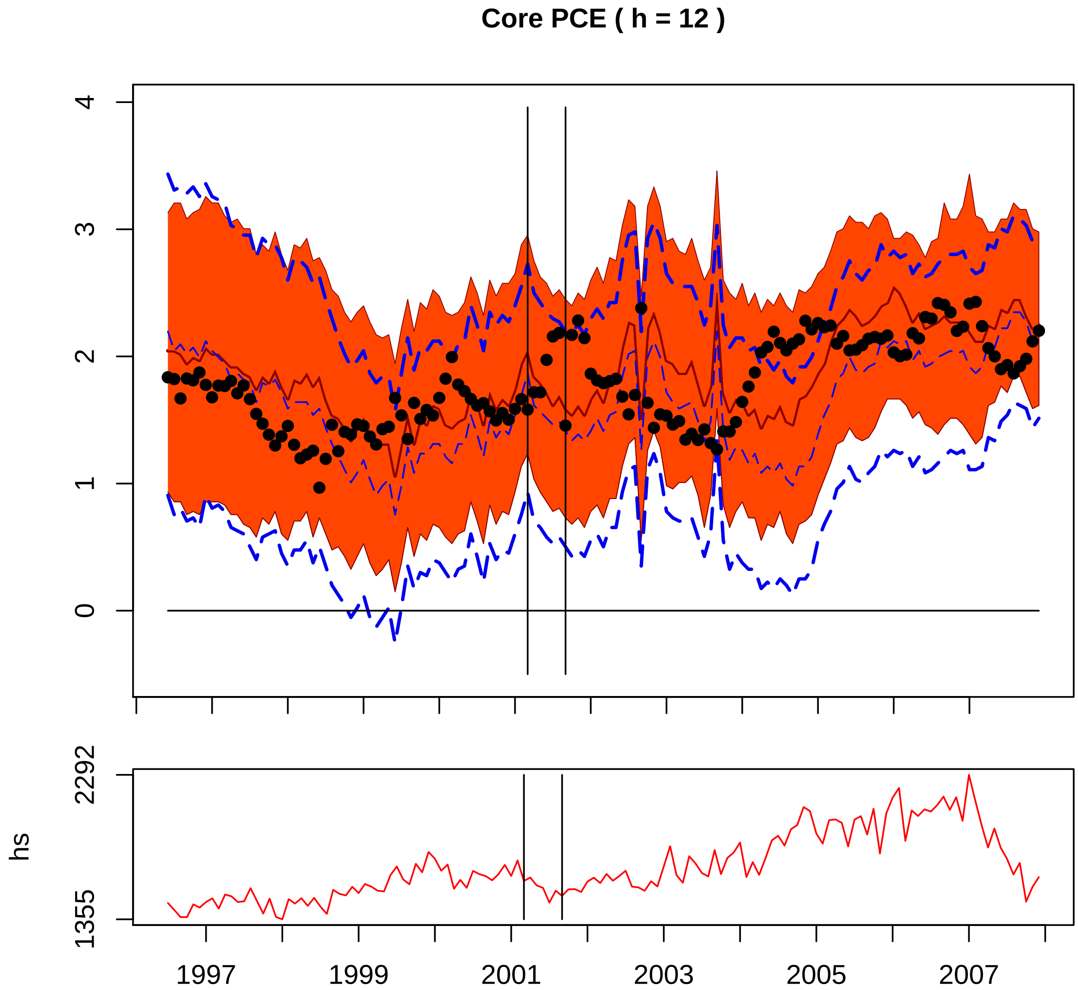


Figure 4: **(Top plot)** The (red) continuous lines denote the quantiles of AO-$U|X^{HS}$ for core-PCE at 5, 50, and 95% levels, the (blue) dashed lines denote the same quantile levels for the AO model, while the circles indicated the realized core-PCE inflation rate. **(Bottom plot)** The housing starts series (shifted forward by $h = 12$ months). The vertical lines indicate the NBER recession dates.

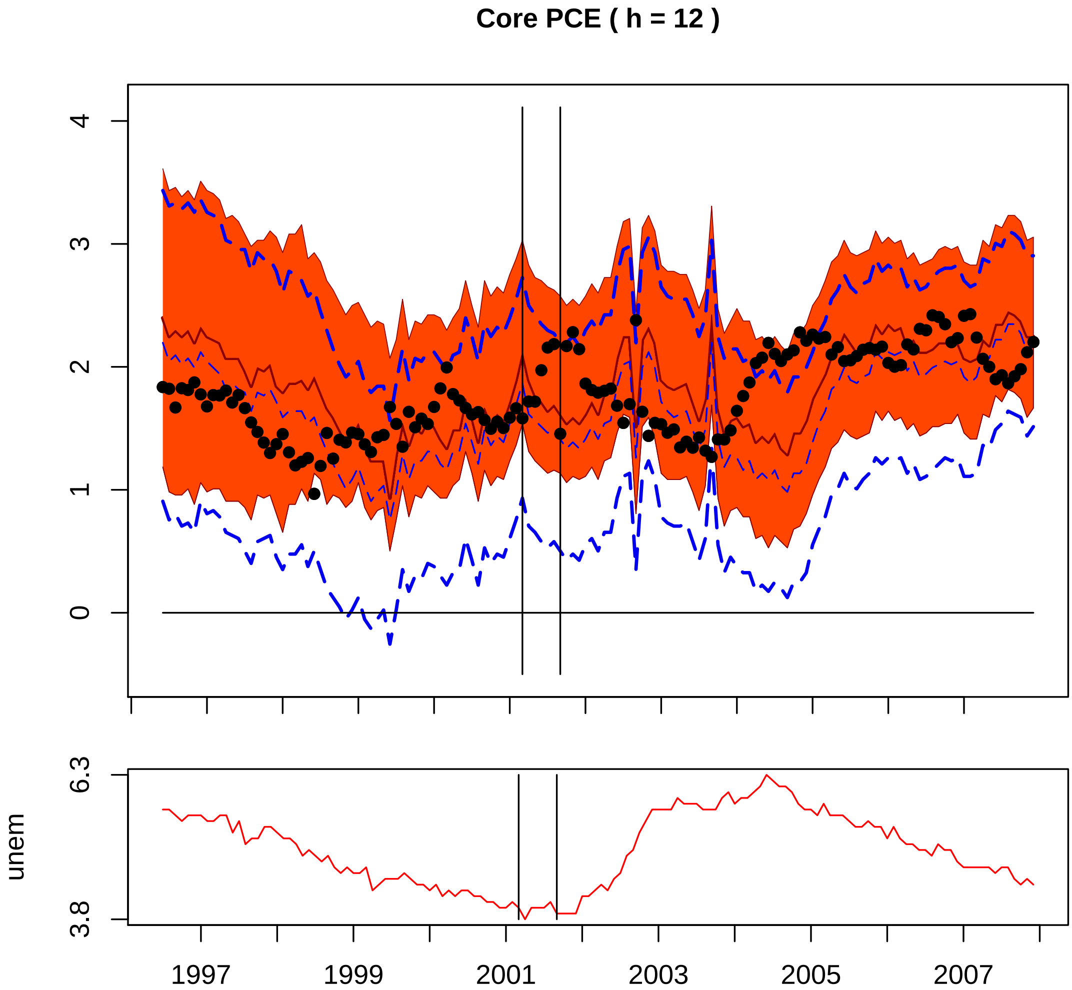


Figure 5: **(Top plot)** The (red) continuous lines denote the quantiles of AO-$U|X^{UNEM}$ for core-PCE at 5, 50, and 95% levels, the (blue) dashed lines denote the same quantile levels for the AO model, while the circles indicated the realized core-PCE inflation rate. **(Bottom plot)** The unemployment rate series (shifted forward by $h = 12$ months). The vertical lines indicate the NBER recession dates.

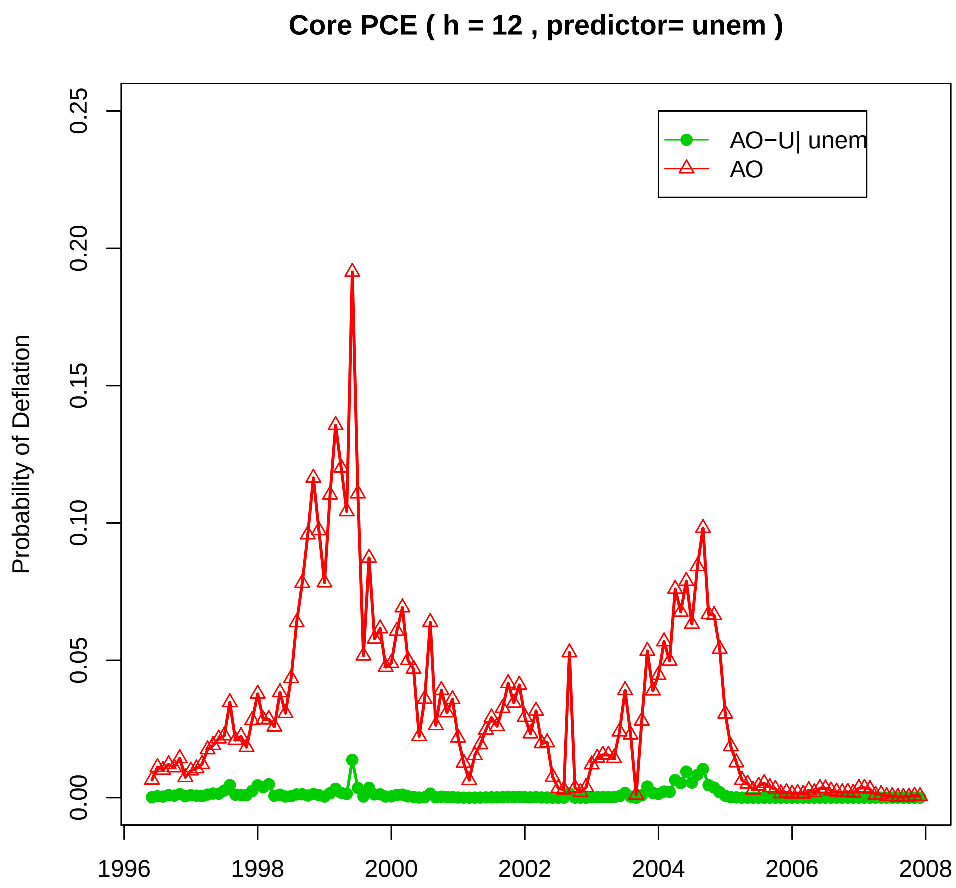


Figure 6: Estimated probability of deflation (defined as a negative inflation rate) for core-PCE using the AO model and AO-$U|X^{UNEM}$ for the period 1996:7-2007:12.