ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Forecasting Under Structural Break Uncertainty

Jing Tian and Heather M. Anderson

July 2011

Working Paper 08/11

Forecasting Under Structural Break Uncertainty

Jing Tian*
University of Tasmania
Hobart, 7005
AUSTRALIA

Heather M. Anderson
Monash University
Clayton, Victoria 3800
AUSTRALIA

Abstract

This paper proposes two new weighting schemes that average forecasts using different estimation windows to account for structural change. We let the weights reflect the probability of each time point being the most-recent break point, and we use the reversed ordered Cusum test statistics to capture this intuition. The second weighting method simply imposes heavier weights on those forecasts that use more recent information. The proposed combination forecasts are evaluated using Monte Carlo techniques, and we compare them with forecasts based on other methods that try to account for structural change, including average forecasts weighted by past forecasting performance and techniques that first estimate a break point and then forecast using the post break data. Simulation results show that our proposed weighting methods often outperform the others in the presence of structural breaks. An empirical application based on a NAIRU Phillips curve model for the United States indicates that it is possible to outperform the random walk forecasting model when we employ forecasting methods that account for break uncertainty.

Keywords: Forecasting with Structural Breaks, Parameter Shifts, Break Uncertainty, Structural Break Tests, Choice of Estimation Sample, Forecast Combinations, The NAIRU Phillips Curve.
JEL classification: C22, C53, E37

---

*Corresponding author. Author contact details: Tian: jing.tian@utas.edu.au and Anderson: heather.anderson@monash.edu. We are grateful for very useful discussions with Farshid Vahid.

1 Introduction

The forecasting of economic variables is often complicated by the possibility that the parameters in the underlying data generating process (DGP) might have changed at various points in time during the pre-forecast sampling period. In this paper, we define structural breaks as permanent shifts in parameters of a DGP, and we focus on the problem that structural breaks often affect forecasts that rely on model estimation. The failure to identify in-sample breaks that change the data generating process produces biased parameter estimates and thus contaminates the model's out-of-sample forecasting performance. Ideally, if information on breaks such as breakpoints and break sizes is known, we can choose the estimation window size according to the trade-off between the bias and forecast error variance to improve the out-of-sample mean squared forecasting errors (see Pesaran and Timmermann, 2007). However, as Pesaran and Timmermann (2007) point out, forecasters usually have little knowledge about structural breaks that might have occurred and therefore, exploiting the bias-variance trade-off becomes difficult in practice.

A conventional forecasting strategy under structural break uncertainty is to select a single estimation window to generate a single forecast. Structural break tests can be applied to test for the presence of breaks, and as a byproduct of most of break tests breakpoints can be estimated if parameter constancy is rejected. Then only data past the estimated most recent breakpoint is included in the estimation window. However, this method excludes all pre-break observations and hence ignores the trade-off between bias and forecast error variance. Moreover, whether breakpoints can be estimated precisely depends on the size of breaks and assumptions relating to the number of breaks or the minimum distance between breaks. If these assumptions are far from the true DGP, then breakpoints might be estimated imprecisely, and a single forecast that relies on imprecisely estimated breakpoints is unlikely to be reliable.

An alternative forecasting strategy that overcomes the shortcomings of a single forecasting strategy is to combine forecasts that use different estimations windows. Pesaran and Timmermann (2007) demonstrate that the MSFE-based weighting method and simple forecast averaging work quite well, even though these methods do not use explicit information about potential breaks and therefore base some of the model estimation on pre-break data. Given that these weighting methods can improve forecasting performance, we conjecture that combination forecasts based on weights that incorporate information on structural breaks might lead to further improvements. Under break uncertainty, a straightforward way to obtain relevant information is to conduct structural break tests.

The main motivation for this paper is to propose alternative combination techniques that improve forecasts under structural break uncertainty. There are two main contributions of this paper. First, we propose a new weighting scheme that utilizes the recursively ordered Cusum squared (ROC) test results to average the forecasts based on different estimation windows. We also propose a simple forecasting weighting technique that effectively just places more weight on forecasts derived from more recent samples. Second, we evaluate the ability of the NAIRU Phillips curve to forecast U.S. 12-month inflation changes using our methods that deal with structural break uncertainty, and we find improved forecasting performance, in contrast to previous work on forecasting inflation.

The idea of combining forecasts across various estimation samples reflects the theoretical point in Pesaran and Timmermann (2007) that forecasting performance can be improved by including pre-break data. We use the ROC test proposed by Pesaran and Timmermann (2002) to obtain knowledge about the most recent breakpoint under break uncertainty, treating each in-sample time as a possible most-recent break and generating forecasts using data after this time. The averaging weights for forecasts are chosen to reflect the probability of each in-sample time being the last break, and we use ROC test statistics and a prior function on the location of breaks to capture this idea. We compare our proposed combination forecasts with other forecasts that allow for break uncertainty, and find that they perform well under a variety of simulated situations.

To investigate the forecasting performance of the proposed techniques that account for structural breaks in practice, we employ the unemployment-based NAIRU Phillips curve model to forecast U.S. 12-month inflation changes. Although Stock and Watson (1999) find statistically significant shifts in the coefficients of the NAIRU model, they claim that existing methods such as rolling regressions that account for parameter instability do not produce better forecasts than expanding window forecasts, and therefore they ignore the breaks. It seems that other authors follow the same reasoning and only focus on full sample estimation (see Atkeson and Ohanian, 2001; Fisher et al., 2002). An important contribution of this paper is to show that it is useful to deal with structural breaks when forecasting inflation.

Our out-of-sample forecasting results show that many types of forecasts that allow for breaks improve the forecasting ability of the unemployment based NAIRU Phillips curve model, and defeat the random walk forecasts of zero inflation changes. Further, our proposed combination forecasts weighted by ROC-statistics and the prior function on the breakpoint location achieve a competitive result with equal weighted forecasts and MSFE weighted forecasts. The simple weighted method on the location of the starting time point of each estimation window outperforms all of the other combination methods.

The outline of this paper is as follows. The next section explains the details of the forecasting methods that account for structural break uncertainty, including the new combination weighting schemes. These methods are firstly examined by Monte Carlo simulations in section 3. We then turn in section 4 to employ these approaches to conduct an out-of-sample forecasting exercise of U.S. 12-month inflation changes based on the NAIRU Phillips curve. Section 5 concludes.

2 Forecasting Methods

Assume that the following linear model is subject to m structural breaks $(T_1, T_2, ..., T_m)$:

$$Y_t = X_t' \beta_{T_{j-1}+1:T_j} + \varepsilon_t, \quad j = 1, 2, ..., m+1 \text{ and } T_{j-1} + 1 \leq t \leq T_j \tag{1}$$

with $T_0 = 0$ and $T_{m+1} = T$. Here Y_t is the dependent variable at time t and X_t is a $p \times 1$ vector of regressors at time t that may contain lags of the dependent variable and lagged explanatory variables. The $p \times 1$ vector $\beta_{T_{j-1}+1:T_j}$ denotes the values of coefficients of regressors in each segment j that starts from $T_{j-1} + 1$ and ends at T_j, and when a structural break occurs, all p coefficients shift permanently until the next breakpoint. For simplicity the vector of regressors X stays the same across all of the segments. Suppose that we have a sample of T observations, and we set the minimum acceptable estimation window size $\underline{w}$ to be at least $2p$. Some forecasting methods require a test sample for evaluating the forecasting performance of models, so we reserve the last $\tilde{w}$ observations when implementing these methods. The one-step ahead forecast of Y_{T+1} conditional on information up to time T is denoted by $\widehat{Y}_{T+1}$. This is computed based on the OLS estimated parameters, $\widehat{\beta}$.

In the following subsections, we firstly introduce the newly proposed combination methods that deal with structural break uncertainty. Then we review some existing combination methods to combine forecasts using various estimation windows, and forecasting techniques that select a single estimation window.

2.1 Combined Forecasts

2.1.1 Average Forecasts Weighted by ROC statistics

We combine forecasts derived from different post-break estimation windows, treating each past time as a possible most-recent breakpoint. If a time point is more likely to be the most-recent breakpoint, we allow its associated forecast to contribute more to the final combined forecast. Therefore, we use a weighting method that reflects the probability of a break at each time point. Under structural break uncertainty, the break tests provide us information about the likelihood that each time location is the most-recent breakpoint in the sense that higher ROC statistics (relative to the relevant critical bounds) will reflect a higher likelihood of a break. Therefore, instead of estimating the distribution of in-sample breaks directly, we use ROC test statistics[1] to construct the weights.

Suppose that the observation matrices $Y_{T:\tau}$ and $X_{T:\tau}$ are in a reversed time order given by:

$$Y'_{T:\tau} = [Y_T, Y_{T-1}, ..., Y_{\tau+1}, Y_\tau], \quad X'_{T:\tau} = [X_T, X_{T-1}, ..., X_{\tau+1}, X_\tau].$$

The location of time τ in the sample $[1 : T]$ is subject to the minimum acceptable estimation window size so that the minimum number of observations subsequent to τ is $\underline{w}$. We derive the least squared estimates of β based on reversed ordered data subsequent to time τ as:

$$\widehat{\beta}^{(R)}_{T:\tau+1} = (X'_{T:\tau+1} X_{T:\tau+1})^{-1} X'_{T:\tau+1} Y_{T:\tau+1}, \quad \tau = T - \underline{w}, T - \underline{w} - 1, ..., 2, 1. \tag{2}$$

The ROC test statistics s_τ are constructed using the squares of standardized one-step-

[1] The use of the reciprocals of p-values from break tests provides an alternative approach for building the weights.

ahead recursive residuals v_t from the reversed ordered regression:

$$s_\tau = \frac{\sum_{t=\tau}^{T-\underline{w}} v_t^2}{\sum_{t=1}^{T-\underline{w}} v_t^2}, \quad \tau = T - \underline{w}, T - \underline{w} - 1, ..., 2, 1, \tag{3}$$

where v_t is computed as:

$$v_t = \frac{Y_t - X_t' \widehat{\beta}_{T:t+1}^{(R)}}{\sqrt{1 + X_t'(X_{T:t+1}' X_{T:t+1})^{-1} X_t}}. \tag{4}$$

Reserving $\underline{w}$ observations, we consider all past dates $\tau \in [1 : T - \underline{w}]$ as a sequence of choices for the last breakpoint, and we approximate the associated probabilities using weights that incorporate the absolute values of the distance between the calculated ROC statistics s_τ and the mid-points of the two ROC critical values given by $\frac{T-\underline{w}-\tau+1}{T-\underline{w}}$. The combination weight on each choice of the most recent breakpoint τ is then constructed using

$$cw_\tau = \frac{|s_\tau - (\frac{T-\underline{w}-\tau+1}{T-\underline{w}})| l_\tau}{\sum_{\tau=1}^{T-\underline{w}} |s_\tau - (\frac{T-\underline{w}-\tau+1}{T-\underline{w}})| l_\tau}, \tag{5}$$

The main intuition behind these $cw_\tau$is that the farther the ROC statistic is away from the mid-point of the two critical values, the more likely it goes across the critical value lines, implying a higher probability of a parameter shift at the associated time point. This combination method provides an averaged forecast given by

$$\widehat{Y}_{T+1} = \sum_{\tau=1}^{T-\underline{w}} cw_\tau (X_{T+1}' \widehat{\beta}_{\tau+1:T}). \tag{6}$$

Note that $\widehat{\beta}_{\tau+1:T}$ is estimated using available observations subsequent to time τ.

The l_τ functions indicate a prior belief on the probability of time τ being the most recent break. The specification of l_τ depends on forecasters' knowledge of structural breaks. For instance, for a short period T when the presence of a single break seems equally likely at each time point, we can define $l_\tau = 1$ for all τ, which makes the weights cw_τ depend only on the magnitude of the ROC statistics. We call the average forecasts obtained by setting $l_\tau = 1$, the ROC-weighted forecasts. However, if we have more historical data and believe that multiple structural breaks may have occurred in the past, this specification becomes less sensible. The reasoning is as follows: If the ROC statistics are observed to be far from the mid-point of the critical-value lines early on in the sample, then with $l_\tau = 1$, the average forecasts will rely heavily on the forecasts generated from the information subsequent to early breakpoints indicated by the ROC statistics. To incorporate the idea that the identification of the most-recent break is more helpful in a forecasting context, we can define the prior weight l_τ to be a function of the location of time τ in the full sample $[1 : T]$, such that

$$l_\tau = \frac{\tau}{T}. \tag{7}$$

The use of (7) in (5) implies that heavier weights are placed on the forecasts based on more recent parts of the sample, because the most-recent break is more likely to happen at the end. By setting $l_\tau = \frac{\tau}{T}$, we reduce the weights on forecasts using information from the beginning of the sample even if the ROC statistics suggest the presence of early structural breaks. In order to distinguish this new weighting scheme from the ROC weights that we have discussed above, we call it an adjusted-ROC weighting scheme, and the associated forecasts are called adjusted-ROC weighted forecasts.

2.1.2 Location Weighted Forecasts

In the forecasting combination literature, many researchers favor equally weighted forecasts since they are simple to compute and often perform better than other elaborate forecasts (see Stock and Watson, 1999; Pesaran and Timmermann, 2007; Clark and McCracken, 2007). In the context of combining forecasts based on different estimation windows, we compute the combined forecasts as:

$$\widehat{Y}_{T+1} = \frac{1}{T - \underline{w}} \sum_{\tau=1}^{T-\underline{w}} (X'_{T+1}\widehat{\beta}_{\tau+1:T}). \tag{8}$$

One can see why this weighting technique might work well when there are breaks by analyzing the trade-off relationship between bias and forecasting error variance, but since individual forecasts use different estimation windows, we can also look at the intuition of the equal weighting method from the perspective of how to weight observations.

Individual forecasts are generated by expanding the length of the estimation window backwards after reserving the most recent $\underline{w}$ observations. Therefore, these $\underline{w}$ most recent observations are used in all of the forecasts, whereas older observations are used less. The use of an equal weighting scheme to average these forecasts essentially modifies the influence of each observation to the combined forecast, so that it has less influence if it is further away from the forecasting origin. The fact that this equal weighting scheme performs well in the presence of structural breaks indicates the importance of the more recent observations when dealing with breaks. It is then quite natural to consider a combination method that relies on the recent data much more than under the equal weighting scheme. A simple method is one that sets the weights to be proportional to the location of time τ in the whole sample $[1 : T]$, i.e τ/T. Note that this weight is the same l_τ that we used when constructing our adjusted-ROC weighted forecasts, and it assigns heavier weights to the forecasts based on more recent samples. The associated combined forecasts, named the location-weighted forecasts, are given by

$$\widehat{Y}_{T+1} = \sum_{\tau=1}^{T-\underline{w}} \frac{\tau/T}{\sum_{\tau=1}^{T-\underline{w}} \tau/T} \left(X'_{T+1}\widehat{\beta}_{\tau+1:T} \right). \tag{9}$$

2.1.3 MSFE Weighted Forecasts

One existing forecasting combination method is to base weights on relative forecasting performance. For example, under the squared error loss function, the weight for a forecast using data $[\tau+1:T]$ is proportional to the inverse of its associated test sample $MSFE_\tau = \tilde{w}^{-1}\sum_{t=T-\tilde{w}}^{T-1}(Y_{t+1} - X'_{t+1}\widehat{\beta}_{\tau+1:t})^2$, which is computed over the window of $\tilde{w}$ periods prior to time T. Then we consider the whole range of values of $\tau \in 1, 2, ..., T - \underline{w} - \tilde{w}$, and compute the forecasts using data subsequent to each value of τ. The weighted average forecast is then given by:

$$\widehat{Y}_{T+1} = \frac{\sum_{\tau=1}^{T-\tilde{w}-\underline{w}} X'_{T+1}\widehat{\beta}_{\tau+1:T}(\frac{1}{MSFE_\tau})}{\sum_{\tau=1}^{T-\tilde{w}-\underline{w}}(\frac{1}{MSFE_\tau})}. \tag{10}$$

2.2 Single Forecasts

Under breakpoint uncertainty, forecasters usually choose to work with a forecasting model that uses data subsequent to the most recent identified break, where breaks might be identified using break detection techniques such as the two-stage reversed ordered Cusum method (Pesaran and Timmermann, 2002) or the Bai-Perron (2003) method. Alternatively, they might work with a forecasting model that uses a sub-sample of data for model estimation, where the sub-sample is chosen so as to minimize a forecasting loss function over a test period in the sample (see Pesaran and Timmermann, 2007). Since we compare the performance of our proposed combination forecasts with single forecasts in sections 3 and 4 of this paper, we briefly outline the latter techniques below.

2.2.1 Two-Stage ROC Method

The two-stage reversed ordered Cusum (ROC) method was first proposed by Pesaran and

Timmermann (2002) and it is based on a standard Cusum squared test for testing and estimating the most recent breakpoint. In the first step, we choose the first time that the ROC test statistic sequence s_τ given in (3) crosses one of the lines of critical values $(T - \underline{w} - \tau + 1)/(T - \underline{w}) \pm c_0$, in which c_0 can be simulated, as shown in Brown et al. (1975) [2].

Conditional on the detection of parameter shifts, the second step of the two-stage ROC method trims all the data prior to the estimated most-recent breakpoint and uses the post-break sample to estimate the forecasting model. If no break is identified by the ROC test, a full-sample estimation is undertaken.

2.2.2 Bai-Perron Method

The Bai-Perron (2003) method consistently estimates the number of breaks and identifies the break locations. It requires assumptions on the maximum number of breaks, denoted by $\bar{m}$ and the minimum distance between two consecutive breaks, denoted by $\underline{h}$. For $m = 1, 2, ..., \bar{m}$, the estimated locations of breaks $\hat{T}_j$ are derived by minimizing the global sum of squared residuals, such that:

$$(\hat{T}_1, .., \hat{T}_m) = argmin_{(T_1,...,T_m)} \sum_{j=1}^{m+1} \sum_{t=T_{j-1}+1}^{T_j} \left(Y_t - X_t' \widehat{\beta}_{T_{j-1}+1:T_j}\right)^2, \tag{11}$$

where $T_j - T_{j-1} \geq \underline{h}$ and the least squares estimates of coefficients $\widehat{\beta}_{T_{j-1}+1:T_j}$ are associated with the sample between $\widehat{T}_{j-1} + 1$ and $\widehat{T}_j$. If we let $m = 0$, then we assume no structural break and use the full sample to estimate the forecasting model.

Bai and Perron (2003) introduce some approaches to determine the number of breaks for $\{m : 0 \leq m \leq \bar{m}\}$. In this paper, we use the Schwartz information criterion (BIC)

[2]This ROC procedure does not produce consistent estimate of breaks since there is always α probability of falsely rejecting the null hypothesis of parameter constancy, where α is the test level.

to select the number of breaks. Once the optimal break number m has been chosen, we compute the forecast using the sample subsequent to the corresponding estimated last breakpoint, or we use the full sample if BIC suggests that no break has occurred.

2.2.3 Cross-Validation Method

The cross-validation method suggested in Pesaran and Timmermann (2007) provides us with a way to choose the "best" estimation window for forecasting without actually estimating breakpoints. Given a test sample of $\tilde{w}$ observations in $[T-\tilde{w}+1:T]$, the optimal estimation window is chosen to start from time τ^* if it minimizes some criterion such as the MSFE over the test sample:

$$\tau^* = argmin_{(1,..,T-\underline{w}-\tilde{w})}\tilde{w}^{-1}\sum_{t=T-\tilde{w}}^{T-1}(Y_{t+1}-X'_{t+1}\widehat{\beta}_{\tau:t})^2. \quad (12)$$

The forecast $\hat{Y}_{T+1}$ is based on estimated coefficients using the sample $[\tau^*:T]$.

3 Monte Carlo Experiments

To evaluate each forecasting method that deals with different levels of break uncertainty, we consider the following bivariate data generating process and conduct Monte-Carlo simulations on

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \mu_{yt} \\ \mu_{xt} \end{pmatrix} + A_t \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix}, \quad (13)$$

where the variance covariance matrix of errors is set to be $\binom{1\,0}{0\,1}$ across the whole sample.

We focus on a two-break DGP rather than a DGP with a single break firstly because this is more realistic in applied situations, and secondly because we want to examine the forecasting methods' ability to account for the most recent break.

In a two-breaks case in which the first and the second breaks occur at $T_1 = p_1 \times T$ and $T_2 = p_2 \times T$ respectively, parameters in the matrix A_t shift permanently after these two breakpoints. The autoregressive parameters are given by

$$A_t = \begin{cases} \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} & t \leq T_1 \\ \begin{pmatrix} a_{11} + d_{11} & a_{12} + d_{12} \\ 0 & a_{22} + d_{22} \end{pmatrix} & T_1 + 1 \leq t \leq T_2 \\ \begin{pmatrix} a_{11} + d_{11} + d_{11}^* & a_{12} + d_{12} + d_{12}^* \\ 0 & a_{22} + d_{22} + d_{22}^* \end{pmatrix} & t \geq T_2 + 1. \end{cases} \tag{14}$$

This DGP follows the simulation setup in Clark and McCracken (2005), where they evaluate the small sample properties of various in-sample predictive ability tests in the presence of structural breaks. Pesaran and Timmermann (2007) also adopt this DGP for comparing different methods for choosing estimation windows when structural breaks occur. This specification allows all of the coefficients as well as only some or one of them to change when a break occurs. The constants in both equations adjust with the shifts of coefficients of regressors to keep the whole DGP stationary.

Suppose that the full sample contains $T = 100$ observations and we are interested in the forecast of y_{101}. The loss is measured by MSFE. After repeating the simulation 5000 times, we evaluate the various forecasting methods that have been discussed in the last section by comparing their MSFEs with the benchmark forecasts based on full sample estimations when structural breaks are ignored. Table 1 records the ratios of the MSFEs relative to the benchmark forecasts. The columns headed *ROC-W1*, *ROC-W2*, *Equal-W*, *L-W* and *MSFE-W* relate to combination forecasts with the ROC weights, adjusted-ROC weights, equal weights, the location-based weights, as well as weights based on test-sample MSFE. In the next three columns headed *2-ROC*, *B-P* and *C-V*, we consider single forecasts in which the estimation window is determined by the two-stage ROC method, the Bai-Perron method and the cross-validation method. We set the minimum size of estimation windows $\underline{w} = 10$ throughout the whole exercise, so that the minimum distance between two breaks $\underline{h}$ in the Bai-Perron method is also 10. This minimum estimation sample size is greater than twice of the number of estimated parameters. Bai and Perron (2003) suggest that a maximum of 5 breaks is sufficient in most empirical work, so we consider up to 5 breaks when we estimate breakpoints using the Bai-Perron method. BIC is used to choose between no break and up to 5 breaks. When forming MSFE-weighted forecasts and cross-validation forecasts, the test sample prior to time T is set to contain $\tilde{w} = 25$ observations.

In the simulations reported in Table 1, both y_t and x_t are persistent, with their autoregressive parameters being $a_{11} = 0.9$ and $a_{22} = 0.9$. The marginal effect of x_{t-1} on y_t is initially one unit. Here we consider various scenarios where the changes of coefficients, denoted by d_{ij} and d^*_{ij} for $i, j = 1, 2$, have different sizes and signs. We assume two true breaks occur, at one quarter and at three quarters of the way through the sample (i.e $p_1 = 0.25$ and $p_2 = 0.75$)[3]. The smallest relative MSFE in each row is highlighted by a double underline, and the second best forecasting method is highlighted by a single line underneath its relative MSFE.

[TABLE 1 ABOUT HERE]

For the first row, the persistency of y_t drops by a large proportion at both breaks. The single forecasts are almost as good as combination forecasts. The location-weighted

[3]We also set the last breakpoint very close to the end of the sample, i.e $p_2 = 0.90$. In most scenarios, applying the methods that account for breaks reduces the relative MSFEs, but the ranking of the methods remains the same. Results can be provided upon request.

forecast, followed by the single forecast using the Bai-Perron method and the adjusted ROC-weighted forecast, is the most accurate with more than 50% gain relative to the forecasts that ignore the breaks. With the same final value of the autoregressive coefficient of y_{t-1}, the second and the third experiments only differ with respect to the size of each break. When the second break is bigger than the first break, the Bai-Perron method successfully identifies the breakpoints and generates the best forecast among all methods. The location-weighted forecast produces the second smallest relative MSFE, and it outperforms the equally weighted forecasts. In the third row where the second break is smaller than the first, the overall forecasting performance is better than the previous scenario, with the location-weighted and adjusted ROC forecasts being the best two.

The fourth and fifth rows report the results for the scenario in which the autoregressive coefficient a_{11} decreases at the first breakpoint and then increases at the second break. When half of the drop is recovered after the second break, the combination forecasts dominate the single forecasts, with the best performance being produced by the simple equal-weighting method. When a_{11} totally recovers after the second break, the single forecast using the Bai-Perron method to estimate breakpoints delivers the smallest MSFE.

The next two rows report the benefits of using forecasting methods for structural break uncertainty when the coefficient of x_{t-1} changes at each break. The gains from considering breaks are considerable, especially when there is a large upwards shift. Among all of the combined forecasts, relative MSFEs from location-weighted and adjusted-ROC weighted forecasts are particularly low, and the single forecast using the Bai-Perron method is the best of all. However, the performance of the Bai-Perron method depends on the size and the direction of the shifts in a_{12}. For the seventh row, where a_{12} declines by a small amount at each break, weighted forecasts dominate the single forecasts.

The forecasting results when no breaks are present in the past are shown in the last row. The two-stage ROC method achieves the relative MSFE closest to 1. This result is expected, since we apply the full sample to estimate forecasting models when no break can be detected by the ROC test. Without breaks, the sequence of the ROC statistics are roughly the same and then the distance between the ROC statistic and the mid point of two boundaries is about the same for every time point. Therefore, the ROC weights are close to equal weights, resulting in similar relative MSFEs for *ROC-W1* and *Equal-W*. Although no break has occurred, the Bai-Perron method detects one break with a high frequency of 88.3%, giving the worst forecasting result of all.

Overall, the results in table 1 suggest the following: The forecast weighted by the adjusted-ROC weights performs well and almost always outperforms the equally weighted forecast. In the presence of structural breaks, the location-based weighting scheme outperforms all the other weighting schemes, including the "superior" simple average and weights based on the ROC structural break test. It is also more accurate than single forecasts in most cases. The only comparable forecasting method is the Bai-Perron method, that seems to perform slightly better than the location-weighting method when the second break is big and easy to detect. Moreover, the adjusted-ROC weighting scheme that reflects a different prior belief regarding each time as the most-recent breakpoint usually produces better forecasts than the ROC-weight that is determined by the ROC statistics alone.

Between the two methods that each require a test sample, the cross-validation method that produces a single forecast generally performs better than a MSFE-weighted forecast. The single forecast using the Bai-Perron method performs worse than the combined forecasts when small breaks or no breaks have occurred. The reasons might be as follows: Firstly, small breaks are difficult to detect and estimate accurately using the Bai-Perron method. Secondly, as suggested in Pesaran and Timmermann (2007), when breaks are small and we are using a squared error loss function, it is not optimal to use only the post-most-recent break data to estimate the forecasting model.

4 Forecasting Inflation

4.1 NAIRU Phillips Curve Models

Inflation forecasts have important implications for monetary policy makers. Among various models for inflation, a Phillips curve which connects the top two domestic economic burdens, unemployment and inflation, attracts the most attention. Early versions of the Phillips curve implied a durable tradeoff between unemployment and inflation, but nowadays more and more economists advocate a "natural rate" of inflation that guides the economy back to equilibrium (Tobin, 1972). For instance, a specification of the Phillips curve called NAIRU (non-accelerating inflation rate of unemployment) is based on the idea that inflation will increase if unemployment stays below its natural rate. A textbook version of the NAIRU model for 12-month ahead inflation changes is

$$E_t(\pi_{t+12} - \pi_t) = \beta \times (u_t - \bar{u}) = -\beta\bar{u} + \beta u_t, \tag{15}$$

where π_t and u_t denote the inflation and unemployment rates respectively. The NAIRU in (15) is the time invariant $\bar{u}$. This model not only provides researchers with a method to estimate the baseline unemployment rate, i.e. the NAIRU, but it also specifies a popular inflation forecasting model because of its simplicity and backward-looking specification.

The usefulness of the NAIRU Phillips curve for inflation forecasts has been discussed in several papers. Under a simulated out-of-sample framework, Stock and Watson (1999) compare forecasts of U.S. inflation rates at the 12-month horizon from 1970 to 1996 using a variety of the NAIRU Phillips curve-based models. The basic 12-month ahead forecasting models used in their paper is generalized as

$$\pi_{t+12}^{12} - \pi_t = \alpha + \beta(L)x_t + \gamma(L)\Delta\pi_t + \varepsilon_{t+12}, \tag{16}$$

where π_t^{12} is 12-month inflation at time t, and π_t is monthly inflation expressed as an annual rate. The variable x_t is often the unemployment rate, another macroeconomic variable, or a diffuse index measuring aggregate real activity at time t. Stock and Watson find that the conventional model with an unemployment rate gap produces no better forecasts than models based on other measures of real aggregate activity.

Atkeson and Ohanian (2001) provide evidence that 12 month-ahead U.S. inflation forecasts from 1985 to 2001 based on NAIRU Phillips curve models are no better than "flipping a coin". The benchmark forecasting model, which subsequent literature often calls the Atkeson and Ohanian model, predicts no change in the 12 month ahead inflation rate and it is given by

$$E_t\left(\pi_{t+12}^{12}\right) = \pi_t^{12}. \tag{17}$$

In order to make the NAIRU Phillips-curve-based inflation forecasts directly comparable with the benchmark, they revise Stock and Watson's model to

$$\pi_{t+12}^{12} - \pi_t^{12} = \alpha + \beta(L)x_t + \gamma(L)\Delta\pi_t + \varepsilon_{t+12}. \tag{18}$$

Triggered by the debate on the usefulness of the NAIRU Phillips curve, Fisher et al. (2002) examine U.S. inflation forecasts generated from equation (18) for three distinct sample periods during which inflation changes have exhibited different volatility. Their results agree with those from Atkeson and Ohanian (2001) only for the low volatility period. Further, once they change the measurement of inflation or revise the models for a

24-month-ahead forecast horizon, the NAIRU Phillips curve models become favorable.

4.2 Instability of NAIRU Phillips Curve

There is a literature on the question of whether NAIRU Phillips curves can forecast inflation, and much of this focuses on model instability. Based on the well-known specification of autoregressive distributed lagged models for NAIRU Phillips curves, the analysis of parameter stability in previous studies includes the statistical relationship between inflation changes and unemployment rates (or other variables revealing real activity), the persistence of inflation changes, and the NAIRU level that is closely related to the intercept of the model[4].

Atkeson and Ohanian (2001) argue that the relationship between the current unemployment rate and future inflation should vary when the economic environment changes because individuals often adjust their expectations regarding economic variables when policy changes. By simply plotting changes in inflation against current unemployment rates from 1960 to 1999, they observe a flatter negative slope after the mid 80's, meaning a weaker relationship between inflation changes and unemployment rates. The paper by Stock and Watson (1999) supports this, with a series of structural break tests for the presence of a single break showing strong evidence of instability on the coefficients of lagged inflation rates, but not on the unemployment rate coefficients.

Although the autoregressive coefficients of Stock and Watson's models are unstable, Stock and Watson (1999) find that the shifts are quantitatively small and thus they ignore coefficient instability in their forecasts. In (2007), Stock and Watson revisit U.S. inflation forecasts by scrutinizing a univariate inflation process. They suggest that a failure to

[4]Note that as we use the means of squared forecasting errors to evaluate forecasts, any variance change should also result in model instability. However, this type of structural break is beyond the scope of this paper.

vary the autoregressive coefficients may lead to the breakdown of recursive autoregressive distributed lagged inflation forecasts.

The question of whether the NAIRU itself has changed over time has also attracted policymaker and academics' attention. Staiger et al. (1997) model the U.S. NAIRU using a cubic spline and find statistical evidence of a declining shift from the 1980's to the 1990's. To estimate the movement of the NAIRU, Gordon (1997) treats the NAIRU as a time-varying variable that follows a stochastic process. His results confirm a lower NAIRU at the end of the 1990's. However, Staiger et al. (1997) also report that their NAIRU estimates are very imprecise, and thus the forecasts based on different estimates of the NAIRU are essentially the same.

The previous literature on the instability of the NAIRU Phillips curve illustrates the nature of break uncertainty in the parameters of this model. Given this background, we implement different forecasting methods that account for break uncertainty to reexamine the inflation forecasting ability of the unemployment-based NAIRU Phillips curve model.

4.3 The NAIRU Forecasts of the U.S. Inflation

4.3.1 Empirical Model and Data

The main forecasting model (shown in equation (19)) in this paper is a variation of equation (18), in that it incorporates π_t^{12} as an explanatory variable, rather than π_t itself. This model enables a direct comparison of our forecasts that allow for structural breaks with those from the benchmark given in (17).

$$\pi_{t+12}^{12} - \pi_t^{12} = \alpha + \beta(L)x_t + \gamma(L)\Delta\pi_t^{12} + \varepsilon_{t+12}. \tag{19}$$

We measure annual U.S. inflation at time t by computing the 12-month changes of the U.S. core CPI (CPI less food and energy) given by $\pi_t^{12} = 100 \times (\ln P_t - \ln P_{t-12})$. The activity variables x_t in this paper are the unemployment rate, denoted by u_t, or changes of the unemployment rate, denoted by Δu_t.

We use monthly data from 1959:01 to 2007:06 retrieved from DATASTREAM, and we calculate out-of-sample forecasts of 12-month ahead inflation changes for the 10 year period from 1997:07 to 2007:06. Thus the initial estimation starts by employing the information from 1959:01 to 1997:06 for forecasting the inflation change in 1997:07. We then recursively estimate the forecasting models once new information is included.

The use of annual inflation changes rather than inflation itself as the dependent variable means that we treat inflation as a non-stationary or I(1) variable. This is consistent with the empirical properties of the sample.[5]

Most literature includes the level of unemployment rates in the NAIRU Phillips curve[6]. However, if the lags of u_t are not stationary while the dependent variable $\pi_{t+12}^{12} - \pi_t^{12}$ is stationary, this imbalance may lead to a lack of explanatory power, ruining the associated forecasting performance. Therefore, we also consider models with lagged annual changes of monthly unemployment rates as a regressor. Note that since $\Delta u_t = u_t - u_{t-12}$ in this setting, the same lag structure in this model incorporates more historical information of unemployment rates than a model that only includes lags of u_t. In the next section, we report out-of-sample forecasting results generated from both forecasting models.

One minor difference between equation (19) and equation (18) is the way in which inflation changes are measured. Following Stock and Watson's model, Atkeson and Ohanian (2001) define monthly inflation at annual rates as $\pi_t = 1200 \times (\ln P_t - \ln P_{t-1})$ and

[5]Results are available upon request.

[6]See Stock and Watson (1999), Staiger et al. (1997), Atkeson and Ohanian (2001) and Fisher et al. (2002).

use the lags of $\Delta\pi_t$ to predict the 12-month changes of annual inflation $\Delta\pi_{t+12}^{12}$. After examining the movements of both the dependent variable and explanatory variables over time, we find that monthly changes of 1-month inflation at annual rates are extremely noisy, whereas 12-month changes of annual inflation move along a relatively smooth path. Therefore, we replace the noisy regressor in equation (18) with the lagged dependent variable (see equation (19)), resulting in a standard autoregressive distributed lagged model (ADL).

The lag structures of u_t and $\Delta\pi_t^{12}$ are time-variant and are selected by either the Akaike information criterion (AIC) or the Schwartz information criterion (BIC) for each forecast. Generally speaking, BIC penalizes more for a long lag structure than AIC, and therefore the models selected by BIC may reduce coefficient estimation errors. However, when we have a relatively long historical data series or set the minimum estimation window size to be a function of the number of estimated coefficients, then the advantage of BIC is not clear. Therefore, in this empirical forecasting exercise we also use AIC to allow for a long lag structure, and we let the number of lags vary from 1 to 12 for both regressors.

To obtain preliminary knowledge about structural breaks in the evolution of U.S. 12-month changes in annual inflation, we show the time series plot of $\Delta\pi_{t+12}^{12}$ from 1959:01 to 2007:06. Figure 1 displays a dramatic volatility change around 1984[7]. However, since the parameter stability in the NAIRU Phillips curve is the main focus of this paper, we need to conduct structural break tests. Although the year 1984 might be a candidate breakpoint, we are still unsure about the number of breaks and whether December 1983 might be the last structural break that shifts parameters. Therefore, in this paper, we conduct two sets of forecasting exercise for U.S. inflation changes: one assumes a break

[7]The literature often interprets this change in the mid 1980's as a result of a shift in monetary policy. For instance, Clarida et al. (2000) show that the U.S. macroeconomy has been stable since the appointment of Paul Volker as the Fed chairman in 1979.

uncertainty problem throughout the whole given sample, and the other assumes break uncertainty after January 1984 and shortens the sample to include only post-1984 data.

[FIGURE 1 ABOUT HERE]

4.3.2 Results

Figures 2 and 5 summarize the out-of-sample forecasting results through July 1997 to June 2007 when the possible presence of structural breaks is considered. We use bar plots to report the means of squared forecast errors (MSFE) relative to the benchmark forecasts generated from equation (17). Relative MSFEs that are smaller than one imply that there is an advantage in using the NAIRU Phillips curve model over the random walk model for inflation forecasts. Each figure contains two panels. Panel A on the left presents forecasting results when the full sample since 1959:01 is considered, whereas the results when only post-1984 data is considered are presented in Panel B on the right. The first bar on the top in each plot shows the relative MSFE based on expanding-window estimations when possible breaks are totally ignored. Then we report forecasting results for five weighting techniques that combine forecasts using different estimation windows and three break-dealing methods that choose only one estimation window and thus generate a single forecast.

[FIGURE 2 ABOUT HERE]

Figure 2 gives the results based on model (19) with lags of u_t. The lag length is chosen using AIC, BIC or simply set to 1. Panel A reports the results when we examine all given historical information for forecasting annual inflation changes. If we totally ignore the possibility of breaks, we use expanding windows to estimate forecasting models. The relative MSFEs are around twice as big as a forecast of zero change in annual inflation, consistent with the forecasting ability of the unemployment-based NAIRU model found in other literature.

However, the sizes of relative MSFEs drop dramatically after we employ forecasting methods that account for structural break uncertainty. The location-weighted forecasts, denoted by L-W, and the MSFE-weighted forecasts, denoted by MSFE-W, are the best two among the five combined forecasts, regardless of the lag structure. The ROC weighted forecasts (ROC-W1) achieve a relative MSFE close to one and the adjusted-ROC weighted forecasts (ROC-W2) achieve a relative MSFE of less than one when the lag structure is chosen using information criteria. Both ROC-W1 and ROC-W2 outperform the equally weighted forecasts.

The three methods that choose the "optimal" estimation window and produce single forecasts generate better results than combined forecasting methods. The difference between these three methods lies in different recursive estimates of the last break or the "optimal" estimation window. Figure 3 shows the sequence of the recursive estimates of the last break from the Bai-Perron and the two-stage ROC method. Assuming that there can be a maximum of 5 breaks through the full sample, the Bai-Perron method always chooses 5 breaks and recursively estimates the last breakpoint to be October 1983, regardless of the lag specification. In contrast to the Bai-Perron estimates, the two-stage ROC recursive estimation procedure finds that the last breakpoint follows a near monotonic increasing trend as we move towards the last forecasting point, and it picks much later time points as the most-recent breakpoint compared with the Bai-Perron estimates.

[FIGURE 3 ABOUT HERE]

Panel B of figure 2 shows that ignoring the possibility of breaks after January 1984 does not greatly damage the relative forecasting ability of the NAIRU Phillips curve. The relative MSFE from the model selected by AIC is now about 15% smaller than the random walk forecast if we use expanding windows to estimate forecasting models. This is because the end of year in 1983 is one of structural breaks that shift the parameters in the NAIRU Phillips curve model. Recall that in Figure 3 the Bai-Perron method persistently estimates October 1983 as the last breakpoint, and this is why we observe similar relative MSFEs when applying the Bai-Perron method in panel A and the expanding window estimations in panel B.

Despite the fact that the forecasting performance of techniques that do not allow for breaks improves after we only use post-1984 data to estimate forecasting models, dealing with possible structural breaks after 1984 can still improve the forecasting performance of the NAIRU Phillips curve model to some extent. Among the averaged forecasts, the location-weighted forecasts are the best regardless of the lag structure. The simple average as well as our proposed ROC and adjusted-ROC weighting schemes achieve similar forecasting results. The combined forecasts based on historical MSFE do not perform better than the forecasts when possible in-sample breaks are ignored. The most attractive result, that the relative MSFE becomes about 40% smaller than 1, comes from the single forecast based on an ADL(1,1) model when the most recent break is estimated by the Bai-Perron method. Figure 4 plots the related recursive estimates of the most recent break assuming a maximum of 3 breaks in the past[8]. It shows that until the middle of 2002, the last break is estimated to be around the end of 1992. After this, the last break is identified in late 1996 for about 2 years, and then March 2004 is chosen until the end of the forecasting time.

[FIGURE 4 ABOUT HERE]

We redo the forecasting exercise, replacing the level of unemployment rates with the

[8]We set a smaller maximum number of breaks since we have a shorter in-sample period.

12-month changes of unemployment rates. In panel A, figure 5, the results from the BIC specification and ADL(1,1) are the same, which reflects the fact that BIC usually favors one lag for each regressor in the model. Moreover, models with short lag structures now generate better forecasts than models with long lag structure, regardless of forecasting methods. It seems that when we include the lagged u_{t-12} in the model and thus can use more past information for forecasts, a model with a shorter lag length in differenced data is preferred. Further, with a short lag structure, the NAIRU Phillips curves outperform the random walk model when possible breaks are ignored. When models are specified using AIC, the forecasting methods that account for structural breaks reduce relative MSFE quite successfully. The location-weighted forecast that makes the combined forecast heavily rely on the more current information performs the best among the five weighting schemes, and the two-stage ROC method generates the best single forecast.

When using only the post-1984 data, Panel B in figure 5 shows that even if the possibility of the presence of the post-1984 breaks is ignored, the forecasts based on NAIRU Phillips curve defeat random walk forecasts regardless of the lag structure. Although under AIC, considering break uncertainty does not increase the forecasting ability much, with a shorter lag structure, accounting for the break possibility can make the NAIRU Phillips curve model outperform the random walk model much more than when such a possibility is ignored. For example, with a BIC specification, the combination forecasting methods increase the forecasting accuracy of the NAIRU Phillips curve by at least 20% comparing with the expanding window forecasts. In particular, the location-weighted forecasts achieve almost 40% reduction in relative MSFE.

[FIGURE 5 ABOUT HERE]

We conclude the following from the empirical NAIRU Phillips curves. First, unlike the small effect from incorporating model instability that Stock and Watson (1999) claim in forecasting inflation changes, we find that accounting for break occurrence in the past can improve the forecasting ability of the unemployment-based NAIRU Phillips curve model. Second, it appears that coefficient instability is one of the reasons why the random walk model outperforms the unemployment-based NAIRU Phillips curve model. After we solve the in-sample structural break problem, the NAIRU forecasts achieve smaller MSFEs than the random walk forecasts. Moreover, the ROC and adjusted-ROC weighting schemes are able to reduce the MSFEs, and their associated forecasts are comparable with the MSFE-weighted forecasts and the "superior" equally weighted forecasts. The adjusted-ROC averaged forecasts are systematically more accurate than the ROC-weighted forecasts since they benefit from the prior belief that more current information increases forecastability. The location-based weighting scheme outperforms all the other sophisticated weighting schemes determined by historical forecasting performance or by the ROC structural break tests.

5 Conclusion

Financial and macroeconomic time series are often found to be subject to parameter instability, reflecting policy changes or regime switches. Ignoring the presence of breaks may produce biased forecasts that are based on biased parameter estimates. The timing of parameter shifts in the past sample is unlikely to be known to forecasters. Although a large number of papers provide techniques for testing for structural breaks, far less discuss the use of test results in the context of forecasting. If a break test rejects parameter constancy, forecasters may simply estimate the forecasting model using the data subsequent to the last estimated break; otherwise, they conduct full sample estimation. Instead of following the traditional approach, this paper has proposed a new forecasting combination method that utilizes break test statistics. Under structural break uncertainty, we consider each in-sample time as a possible most recent breakpoint and estimate forecasting models excluding data prior to this time point. We then use weights that reflect the approximated probability of each time point being the most-recent break date, and these weights are determined by the reversed ordered Cusum squared statistics together with a prior probability of the break location.

The other weighting scheme proposed in this paper simply assumes that the strength of the effect of a possible break at each time point declines in proportion to the location of this time point from the end of the sample. Under this weighting scheme, we essentially place more weight on the forecasts that use more recent information. Our Monte Carlo simulations examine the forecasting performance of the new combination methods, as well as a range of alternative forecasting techniques that account for break uncertainty. Our results support the new weighting schemes based on the ROC statistics, particularly the adjusted-ROC that puts heavier weight on those forecasts based on samples that the ROC statistics indicate subsequent to the most-recent break and include more recent information. Further, the forecasting performance of the location-weighted forecasts is particularly promising.

We provide empirical evidence of the benefits obtained from using the proposed combination methods as well as the other alternative methods to re-examine the inflation forecasting ability of the NAIRU Phillips curve for the U.S.A. After taking possible in-sample structural breaks into account, our inflation forecasts are more accurate than the random walk forecasts that Atkeson and Ohanian (2001) prefer. Our results also indicate the necessity of considering the presence of structural breaks when we forecast U.S. inflation changes using the NAIRU Phillips curve.

It is interesting to see that both our Monte Carlo and empirical experiments show that the location-based combination method forecasts well. This weighting scheme not only beats the other more sophisticated weighting methods that utilize information from structural break tests or from the past forecasting performance, but it also outperforms the simple average method that has been commonly used, and is typically hard to beat. The intuition behind this location weighting scheme is to impose more weight on forecasts that use more recent sample periods, and a detailed investigation on this location-based weighting technique is the subject of future research.

References

Atkeson, A. and L. Ohanian (2001): "Are Phillips Curve Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 25, 2–11.

Bai, J. and P. Perron (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18, 1–22.

Brown, R. L., J. Durbin, and J. M. Evans (1975): "Techniques for Testing the Constancy of Regression Relationships over Time," *Journal of the Royal Statistical Society, Series B (Methodological)*, 37, 149–192.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147–180.

Clark, T. E. and M. W. McCracken (2005): "The Power of Tests of Predictive Ability in the Presence of Structural Breaks," *Journal of Econometrics*, 124, 1–31.

——— (2007): "Averaging Forecasts from VARs with Uncertain Instabilities," FRB of Kansas City Working Paper No. 06-12; FEDS Working Paper No. 2007-42; FRB of St. Louis Working Paper 2008-030A .

Fisher, J. D. M., C. T. Liu, and R. Zhou (2002): "When Can We Forecast Inflation," *Economic Perspectives*, 30–42.

Gordon, R. J. (1997): "The Time-Varying NAIRU and Its Implications for Economic Policy," *The Journal of Economic Perspectives*, 11, 11–32.

Pesaran, M. H. and A. Timmermann (2002): "Market Timing and Return Prediction Under Model Instability," *Journal of Empirical Finance*, 9, 495–510.

——— (2007): "Selection of Estimation Window in the Presence of Breaks," *Journal of Econometrics*, 137, 134–161.

Staiger, D., J. H. Stock, and M. W. Watson (1997): "The NAIRU, Unemployment and Monetary Policy," *The Journal of Economic Perspectives*, 11, 33–49.

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

——— (2007): "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, 39, 3–33.

Tobin, J. (1972): "Inflation and Unemployment," *The American Economic Review*, 62, 1–18.

Figure 1: 12-month changes of U.S. annual core CPI inflation

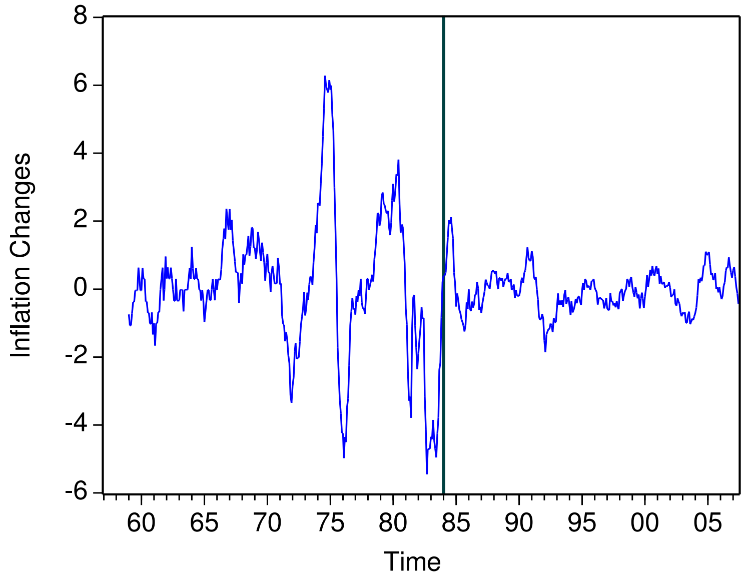

Figure 2: Relative MSFE of Phillips-curve based inflation forecasts (all models use the level of the unemployment rate as one of the regressors)

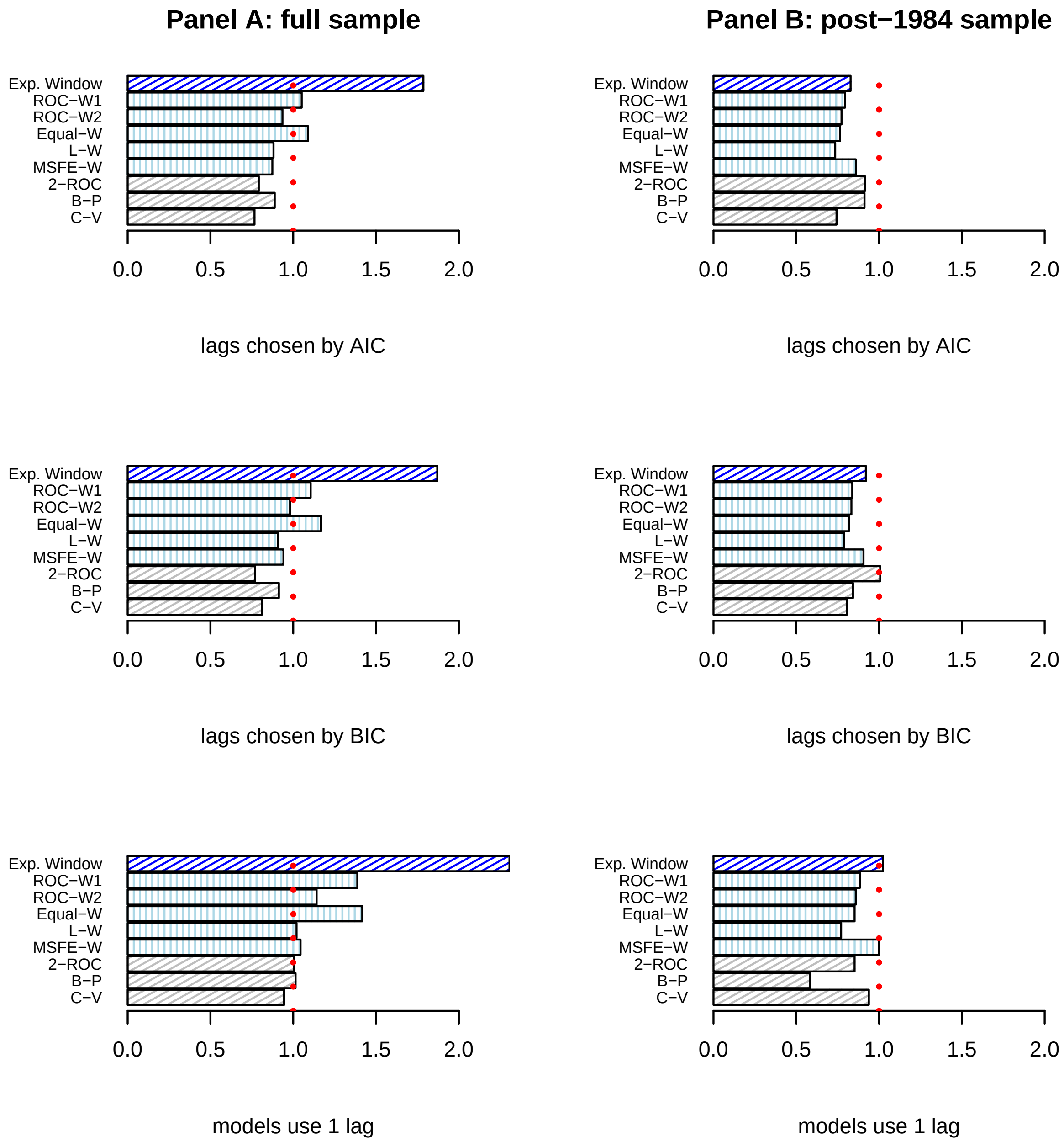

Figure 3: Recursively estimated last breakpoints

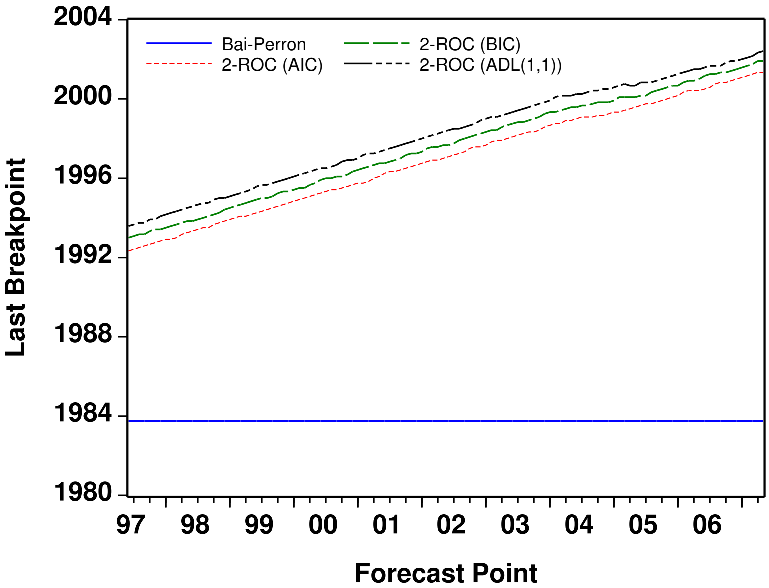

Figure 4: Recursively estimated last breakpoints

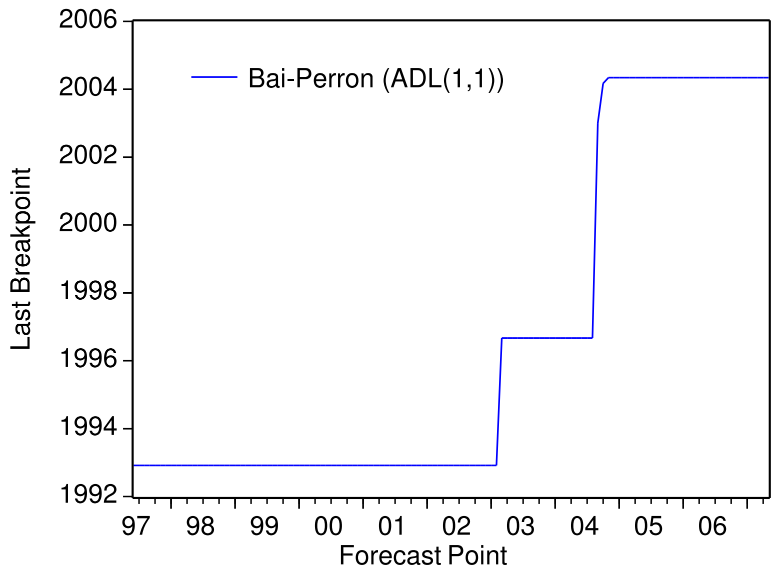

Figure 5: Relative MSFE of Phillips-curve based inflation forecasts (all models use 12-month changes of unemployment rate as one of the regressors)

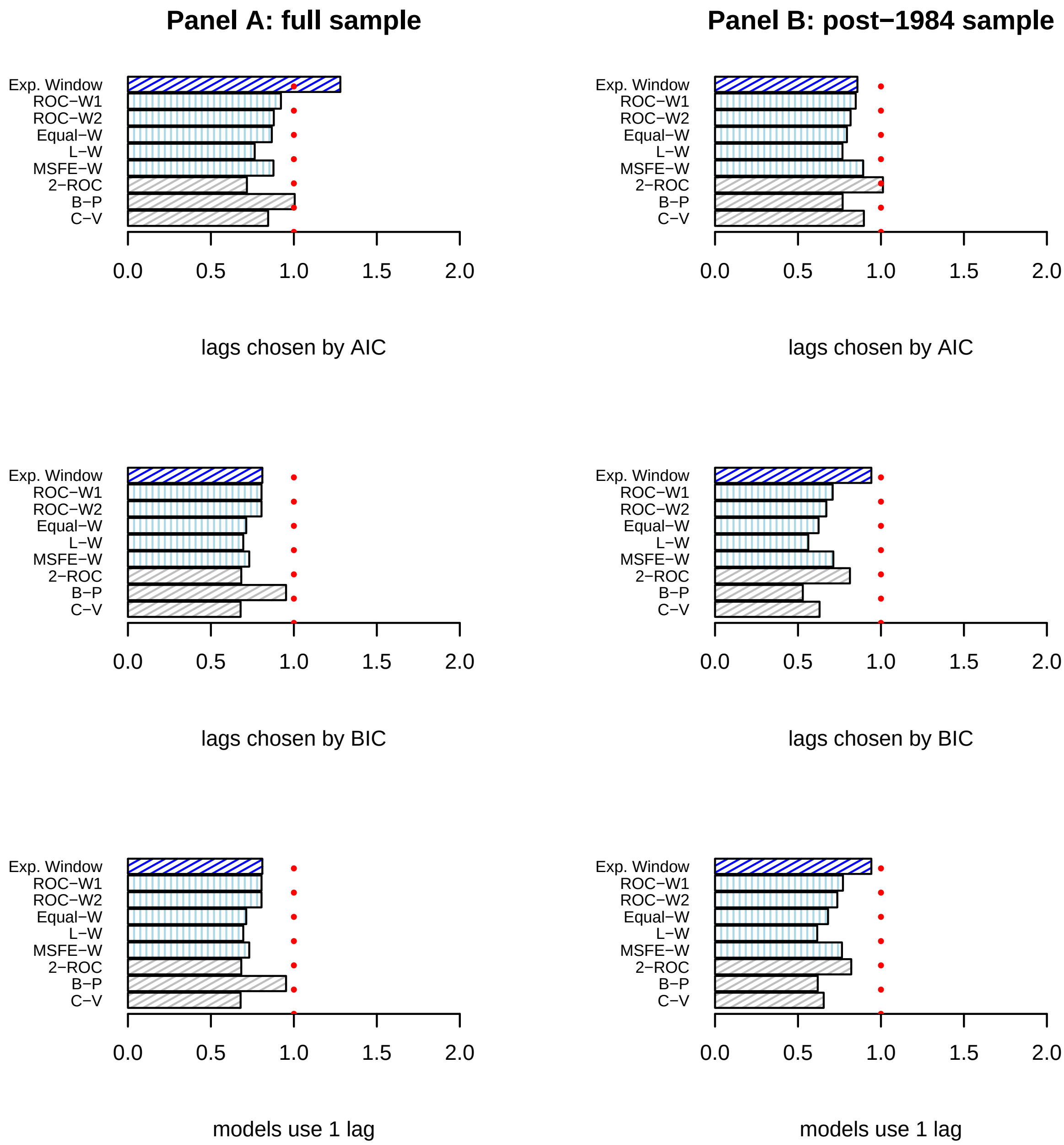

Table 1: Relative MSFE values under multiple structural breaks[a]

							Combined Forecast					Single Forecast		
No.	d_{11}	d_{12}	d_{22}	d^*_{11}	d^*_{12}	d^*_{22}	*ROC-W1*	*ROC-W2*	*Equal-W*	*L-W*	*MSFE-W*[b]	*2-ROC*	*B-P*[c]	*C-V*[b]
1	-0.4	0	0	-0.4	0	0	0.527	0.489	0.516	0.453	0.597	0.539	0.481	0.510
2	-0.2	0	0	-0.4	0	0	0.613	0.552	0.589	0.507	0.713	0.760	0.489	0.612
3	-0.4	0	0	-0.2	0	0	0.513	0.498	0.522	0.490	0.555	0.519	0.553	0.514
4	-0.4	0	0	0.2	0	0	0.799	0.784	0.758	0.761	0.785	0.831	0.903	0.813
5	-0.4	0	0	0.4	0	0	0.867	0.790	0.840	0.753	0.969	0.975	0.742	0.859
6	0	1	0	0	1	0	0.386	0.318	0.371	0.282	0.472	0.506	0.243	0.352
7	0	-0.4	0	0	-0.4	0	0.722	0.699	0.711	0.687	0.775	0.917	0.812	0.748
8	0	0	0	0	0	0	1.030	1.058	1.032	1.075	1.010	1.005	1.362	1.027

[a] All forecast performances are relative to a model that makes no allowance for breaks. We set the values of parameters to be $a_{11} = 0.9, a_{12} = 1, a_{22} = 0.9$ and use $T = 100$ in-sample observations. The true locations of the two structural breaks are $T_1 = p_1 \times T$, where $p_1 = 0.25$ and $T_2 = p_2 \times T$ where $p_2 = 0.75$. We set the minimum estimation window to be $\underline{w} = 10$. Double and single underlines indicate the best and second best forecasting performances.

[b] We set the test sample size $\tilde{w} = 25$.

[c] The number of breaks is chosen from 0 to 5 by BIC.