November 7, 2006

Forecasting Persistent Data with Possible Structural Breaks: Old School and New School Lessons Using OECD Unemployment Rates

Walter Enders, University of Alabama

Ruxandra Prodan, University of Houston

Abstract

In contrast to recent forecasting developments, 'Old School' forecasting techniques, such as exponential smoothing and the Box-Jenkins methodology, do not attempt to explicitly model or to estimate breaks in a time series. Adherents of the 'New School' methodology argue that once breaks are well-estimated, it is possible to control for regime shifts when forecasting. We compare the forecasts of monthly unemployment rates in 10 OECD countries using various Old School and New School methods. Although each method seems to have drawbacks and no one method dominates the others, the Old School methods often outperform the New School methods for forecasting the unemployment rates.

Keywords: Bai-Perron test, Nonlinear forecasting, Out-of-sample Forecasting

JEL classifications: C22, C53

Correspondence to:

Walter Enders*, tel. (205) 348-8972, email: wenders@cba.ua.edu, Department of Economics, University of Alabama, Tuscaloosa, AL 35487.
Ruxandra Prodan, tel. (713) 743-3836, email: rprodan@mail.uh.edu, Department of Economics, University of Houston, Houston, TX 77204.

It is well-known that unemployment rates in the majority of the OECD countries exhibit high persistence and are likely to be characterized by an unknown number of structural breaks. Moreover, there is evidence that these rates display asymmetric adjustment over the course of the business cycle. This complicates the task of forecasting unemployment rates because it is often difficult to distinguish between multiple structural breaks, high persistence, and nonlinearities in the data-generating process.

The recent forecasting literature seems to have shifted its orientation regarding the appropriate way to forecast with structural breaks. ‘Old School’ forecasting techniques, such as exponential smoothing and the Box-Jenkins methodology, do not attempt to explicitly model or estimate the breaks in the series. Instead, forecasters, using some form of exponential smoothing, would account for changes in the level of the series using forecasts that place relatively large weights on the most recent values of the series. Forecasters using the Box-Jenkins methodology, might simply first-difference or second difference the series in order to control for the lack of mean reversion. The level of differencing could be chosen by an examination of the autocorrelation function or by the use of some type of Dickey-Fuller test. In contrast, ‘New School’ forecasters attempt to estimate the number and magnitudes of the breaks. Once the break dates have been estimated, they can be used to control for regime shifts when forecasting.

In this paper, we perform a forecasting competition between various Old School and New School methods of treating structural change. One way to conduct such a ‘horserace’ is to perform a Monte Carlo experiment using simulated data. The advantage of this approach is that the essential characteristics of the data generating process (i.e., the actual number of breaks, the size of the breaks, the location of the breaks, and the persistence) are known to the researcher. Of course, the main disadvantage of such an approach is that the simulated data may not resemble to the actual types of economic data used by forecasters. Moreover, someone considering which method to use for forecasting a particular series may have little knowledge of the actual nature of the breaks. As such, we perform our forecasting competition using monthly unemployment rates in 10 OECD countries. As discussed in more detail below, these series seem to contain structural breaks and all are highly persistent. The paper is organized as follows. Section 1 discusses some of the issues concerning forecasting a persistent series with potential structural breaks. Section 2 examines some of the time series properties of the monthly unemployment rates in 10 OECD countries. All the series are highly persistent and a cursory examination indicates that they are all likely candidates to have at least one break. Section 3 discusses the specific methodologies we use to construct the expanding window forecasts. Section 4 contains our key results. We show that each method seems to have drawbacks and no one method dominates the others. Surprisingly, for this data set, the Old School methods usually perform far better than the New School methods. Concluding remarks are contained in Section 5.

1. Forecasting with Structural Breaks

Clements and Hendry (1999) argue that a prime source of forecast failure is improperly modeled structural change. Furthermore, Clark and McCracken (2005) argue the that large differences between a model's in-sample fit and out-of-sample forecasts might be due to the parameter instability resulting from structural breaks. Many of the issues involved when forecasting in the presence of structural breaks are illustrated by the simulated time series variable y_t shown in Panel *a* of Figure 1. The dashed line represents the quarterly observations of the following simulated AR(1) process with an autoregressive coefficient of 0.8 and two breaks in the mean:

$$y_t = \alpha(t) + 0.8(y_{t\text{-}1} - \alpha(t)) + \varepsilon_t \tag{1}$$

where: $\alpha(t) = 1$ for 1968:1 < t ; $\alpha(t) = 4$ for 1968:1 ≤ t < 1990:1 ; $\alpha(t) = 2.5$ for 1990:1 ≤ t; and ε_t is ~ $N(0, 1)$. The solid line in the figure represents the time path of the unconditional mean $\alpha(t)$. Note that the span of the data shown in the figure roughly corresponds to the actual time period of the unemployment rate series we use below.[1]

Since the last value of y_t equals 3.22 (i.e., $y_{2005:4} = 3.22$), the optimal forecasts for subsequent values of the series can be obtained from (1) by setting $\alpha(t) = 2.5$, $y_{t-1} = 3.22$ and calculating the recursive forecasts while assuming that all future $\varepsilon_t = 0$. Hence, the optimal 1-step, 2-step, and 3-step ahead forecasts are $y_{2006:1} = 3.08$, $y_{2006:2} = 2.96$ and $y_{2006:3} = 2.87$. Clearly, the forecasts converge to the unconditional mean in such a way that 80% of each discrepancy between the forecast and the mean value of 2.5 persists into the next period.

Of course, in applied work, the researcher does not know the magnitude of the autoregressive coefficient or the time path of $\alpha(t)$. Having observed only the dashed line representing y_t, how can reasonable forecasts be constructed? Simply ignoring the breaks and estimating the entire series as an AR(1) process has the downside that the estimate of the persistence parameter is likely to be biased and the forecasts will not converge to the actual value of $\alpha(t)$.

The Old School methods of dealing with a series that does not appear to be mean reverting include exponential smoothing and differencing. Panel *b* of Figure 1 shows the actual values of y_t and the set of 1-step ahead forecasts (f_t^1) calculated from the simple smoothing function:

---

[1] Our notation is such that *year*:*p* denotes the *p*-th period of the specified year. Hence, with quarterly data, 1990:1 is the first quarter of 1990. The ten unemployment rate series we use to forecast are all monthly. Hence, for monthly data, 1990:1 would refer to January 1990.

$$f_t^1 = f_{t-1}^1 + \beta(y_{t-1} - f_{t-1}^1) \quad (2)$$

where: at each period t, the value of β was chosen to minimize the sum of the in-sample square prediction errors, and the dashed line in the panel depicts the 1-step ahead forecasts.

The point illustrated by the figure is that it is hard to discern the actual values of y_t from the forecasted values. Since the exponential forecasts place a relatively large weight on the most recent values of the series (over the sample period, the estimated value of β averaged about 0.85), the methodology acts to quickly capture the effects of a shift in the mean of the y_t series.

Users of the Box-Jenkins method would begin by examining the autocorrelation function (ACF) of y_t. As shown in Panel *c* of Figure 1, the ACF for y_t slowly decays to zero and then begins to increase. As such, the standard practice would be to estimate the series in first-differences so as to remove the effects of the level-shifts from the series. Panel *d* of the Figure 1 shows the 1-step ahead forecasts for y_t obtained by estimating the Δy_t series as an AR(1) process. Again, since the actual series and the 1-step ahead forecasts are very similar to each other, it should be clear that differencing allows a researcher to obtain reasonable forecasts of a series with several level shifts. Sometimes second-differencing is used when breaks are suspected. For example, Clements and Hendry (1999) show that second-differencing the variable of interest improves the forecasting performance of autoregressive models in the presence of structural breaks.

Many extensions of the Box-Jenkins methodology have been proposed. Instead of simply differencing a series based on its ACF, Diebold and Kilian (2000) argue that pre-testing for a unit root using a Dickey-Fuller test can improve forecast accuracy. The basic idea is to difference the series only if the null hypothesis of a unit root cannot be rejected. However, the benefit of pre-testing is not universally accepted because of the well-known inability of unit-root tests to distinguish a unit root null from nearby stationary alternatives. Similarly, Clements and Hendry (1999) argue that the difference-stationary and trend-stationary models are indistinguishable in terms of their implications for forecastability.

Another common practice is to use “rolling” regressions if structural breaks are suspected. Instead of using all of the data to create sequential forecasts, rolling regressions use a fixed number of observations. As a new observation becomes available, the previous initial observation is dropped from the regression equation. In terms of Figure 1, since a rolling regression excludes the earliest observations, it acts to eliminate some of the observations contained the pre-break data (i.e., observations from Regimes 1 and 2). Nevertheless, this method has numerous disadvantages. Rolling regressions cannot be fully efficient if only some of the parameters change since none of the pre-break data is used to estimate the unchanging parameters. Also note that the methodology is such that an observation receiving full weight for forecasting in one period is relegated to having a weight of zero in the next.

Clements and Hendry (1999) champion the use of an intercept correction (IC) to account for the possibility of structural breaks. An IC adjusts an equation’s constant term when forecasting; specifically, in a model that has been under-forecasting (over-forecasting), the intercept term can be adjusted upward (downward) by the extent of the bias. Of course, the use of an IC is not costless since it can increase the mean square prediction error (MSPE) if the actual data generating process is properly modeled.

The New School strategy is to estimate the timing of the breaks and to take this information into account when forecasting. In terms of Figure 1, if the break dates (1968:1 and 1990:1) were known, it would be possible to estimate and forecast the y_t series using an equation of the form:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \varepsilon_t \quad (3)$$

where: D_{1t} and D_{2t} are Heaviside indicators such that $D_{1t} = 1$ only for $1968{:}1 \leq t < 1990{:}1$ and $D_{2t} = 1$ only for $1990{:}1 \leq t$.

A regression in the form of (3) would yield an estimate for α_1 and the values of γ_1 and γ_2 would yield the estimated magnitudes of the breaks. The advantage of this method is that all of the data could be used to estimate the persistence parameter α_1. In principle, this seems to be a far better approach than the Old School methodologies since all of the data is used and the model used to forecast coincides with the actual data-generating process. However, in most of the cases, the break dates (if any) are unknown and need to be estimated. As such, it seems reasonable to use the Andrews and Ploberger (AP) (1994) optimal break test for a series suspected of containing a single structural break. If the test indicates the presence of a break at date t, a dummy variable, D_t, can easily be created and included in the forecasting model. Moreover, since unemployment rates may experience multiple breaks, it is also reasonable to use the Bai and Perron (BP) methodology to estimate the break dates. The BP methodology generalizes the AP test in that it is designed to estimate multiple structural breaks occurring at unknown dates.

Nevertheless, there are potential drawbacks of using such New School tests for structural change. Clearly, if the break dates and/or the magnitudes of the γ_i are poorly estimated, the out-of-sample forecasts will suffer. This is particularly important since Diebold and Chenn (1996) provide evidence of size distortions in the AP supremum tests for a structural change in a dynamic model. Since the null hypothesis maintains that there is no break in the data generating process, improperly rejecting the null implies the inclusion of spurious breaks in the forecasting equation. Similarly, Prodan (2005) shows that the BP test contains serious size distortions if the data is highly persistent since it is difficult to distinguish between a break and persistence.

Bai and Perron (2001, 2003) acknowledge the lack of power of their sequential procedure on data that includes certain configuration of changes (especially breaks of opposite sign). However, the lack of power of the test can mean that breaks occurring in the data are omitted from the forecasting equation. Moreover, unlike equation (3), the breaks can manifest themselves in the autoregressive coefficients as well as the intercept term. Although the most general way to proceed is to allow for breaks in all of the coefficients, this method loses power if the break affects only a few of the coefficients. Another problem is that the AP and BP tests assume that the data is regime-wise stationary. Given that the regimes themselves are unknown, it is unclear how a researcher knows whether the series is stationary within each regime. Finally, these New School methods require full specification of the dynamics (or the use of consistent estimates of serial-correlation parameters); a poor handling of any of these leading to undesirable properties in finite samples.

Note that the multiple breaks allowed in the BP test can be estimated sequentially or globally (i.e., simultaneously). It is quite possible that alternative methods indicate a different number of breaks. Moreover, once the break dates are estimated, it is unclear whether all of the data should be used for forecasting (i.e., it is unclear whether the data from Regimes 1 and 2 should be used for forecasting past the Post-Break period). In fact, Timmermann and Pesaran (2004) show that forecasting performance is dependent on the correct selection of the post-break window used to forecast. If the break is large, the post-break window forecasting method is more efficient than an expanding or rolling window. On the other hand, Timmermann and Pesaran (2005) find both, theoretically and empirically, that the inclusion of some pre-break data for the purpose of estimation of the autoregressive models parameters' leads to lower biases and lower mean squared forecast errors than if only post-break data is used.[2]

To complicate the issue, many researchers have reported evidence of nonlinearities in various unemployment rate series. Specifically, over the course of the business cycle, increases in unemployment rates tend to be sharp and persistent while decreases in the rate occur slowly. As such, Rothman (1996) has shown that several nonlinear models dominate the linear model when forecasting the out-of-sample US quarterly unemployment rates.[3] For our purposes, a regime-switching model is especially interesting since the switch in the coefficients can be interpreted as a break. One key difference between regime-switching and pure break models is that there is only a small number of values that the coefficients can assume in a switching model.

The aim of the paper is to compare the performance of the Old School and New School methods of forecasting unemployment rates in the OECD countries. By comparing the out-of-sample properties of the various methodologies, we should be able to shed light on the method an applied researcher might want to use for forecasting a persistent series with multiple breaks. Specifically, for each OECD country in our sample, we will estimate the unemployment rate in levels, in first- and second-differences, exponentially smoothed, with and without intercept corrections. We will also determine whether pre-testing for unit roots can lead to any improvement in the forecast accuracy. All of these Old School methods are compared to forecasts using the AP test and various forms of the BP test. We pay particular attention to the various forms of the BP test. In particular, we want to determine whether the break dates should be estimated sequentially or simultaneously (i.e., globally) and whether it is better to estimate the

[2] In the paper we report forecasting results using an expanding window and using only post-break data. In the NS methods, we find little difference between these two methods. As such, we do not report results using a rolling window.

[3] The results are sensitive to whether a first-differencing is applied to the unemployment rates series.

model using only post-break data or all of the data. We consider each of these variants using both parametric and nonparametric methods to estimate the variance. Finally, we compare all of the forecasts of a class of nonlinear time-series models designed to capture asymmetric adjustment.

2. Time-Series Properties of the Unemployment Rates

The notion that the rate of unemployment should fluctuate around a long-run steady state or “natural rate” is one of the central theories of macroeconomics. Within this framework, deviations in unemployment from the natural rate should be temporary implying that the unemployment rate should be stationary. However, the high persistence of unemployment rates in many countries seems to question the validity of the natural rate paradigm. As a result, Blanchard and Summers (1987) theorize that unemployment might be characterized by “hysteresis” so that changes in unemployment rates tend to be permanent in nature. As such, unemployment rates are likely to be unit-root processes. A third theory describing unemployment, [see Phelps (1994)] is that most shocks to unemployment are temporary, but occasionally the natural rate will permanently change. As a consequence, the unemployment rate may be characterized as a process that is stationary around a small number of (permanent) “structural breaks.” The fourth hypothesis, the business cycle asymmetry hypothesis, is that unemployment rates increase quickly in recessions but decline relatively slowly during expansions. As such, a nonlinear time series model might be better suited to capture the dynamics of the unemployment rate than a linear model. All of this ambiguity concerning the nature of the unemployment rates makes forecasting especially difficult since the actual data generating process might be stationary or nonstationary, might include structural breaks, and can be linear or nonlinear.

In order to gain some insight into the time-series behavior of unemployment rates, we examine the out-of-sample forecasting properties Old School and New School methods for monthly unemployment rates in ten OECD countries. The countries that we use in our forecasting experiment are US, Canada, Australia, Netherlands, Norway, Denmark, France, Germany, Japan and UK. The length of the data used varies from country to country: it starts from 1948:1 for the longest of the series (US) and from 1978:02 for the shortest one (Australia).[4] For each country, the starting date of the series in given in the second column of Table 1. All series end in June of 2005 (2005:6).

A visual inspection of Figures 2a, 2b, and 2c suggests that most unemployment rate series exhibit a structural break. Moreover, some of the series show a possible trend or time varying trend (France, Japan, Norway), possible unit roots (Netherlands, Denmark), U-shape (Germany and UK), possible stationarity (US), and other types of nonlinearities. The descriptive statistics reported in Table 1 serve to reinforce these visual impressions. For each series, the sample mean and the 12th-lagged autocorrelation coefficient is reported in column 3 and colums 7, respectively. Since most of the series have a different number of observations, we divided the sample period of each country into thirds. Columns 4 – 6 of the Table 1 report the subsample means and columns 8 – 10 report the subsample 12th-lagged autocorrelation coefficients for each third. For example, the overall mean of the US unemployment rate was 5.63% for the full sample period 1948:1 – 2005:6. However, the rate averaged 4.82% over the 1948:1 – 1967:2 period (the first third of the sample), 6.44% over the second third of the sample and 5.63% over the last third of the sample. The autocorrelation for lag 12 showed little persistence for the first third of the sample period but was approximately 75% for the last two thirds of the sample.

[4] The U.S. rates were obtained from the Bureau of the Labor Statistics. The Canadian, Australian and Netherlands rates were obtained from the OECD and rates for all the other countries were obtained from Global Insight. All of the series are seasonally adjusted.

Even though the overall impression from Table 1 is that there were changes in the means and autocorrelations of most of the series, it is difficult to know how to model these changes. Clearly, dividing each sample into thirds was arbitrary—other divisions would certainly result in different values. The point is that the unemployment rates are highly persistent and are likely to have experienced structural breaks. This makes testing and modeling difficult since it is often difficult to distinguish between structural change and high persistence. In the extreme case of a random-walk, the complete persistence of each shock is tantamount to a structural break in each period.

3. Specifics of the Estimated Models

In this section we will briefly discuss econometric methods used in our analysis of the unemployment rates.

3.1. Old School Models

Our benchmark forecasting model is the linear autoregressive model:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \varepsilon_t \tag{4}$$

The main econometric problem is to determine the lag length p and then estimate the parameters. We choose the number of AR coefficients (p) by minimizing the Schwartz Bayesian Information Criteria (BIC).

In order to consider the possibility of stochastic trends, we also estimate (4) using first and second differences of the unemployment rates. In all cases, we choose the number of autoregressive lags (p) by minimizing the BIC. Next, we follow the suggestion of Diebold and Kilian (2000) and pretest for a unit root using an augmented Dickey-Fuller test. We then forecast in levels or in first-differences depending on the outcome of the pre-test. The issue is to help determine whether pre-testing can improve the out-of-sample forecasting performance.

There are many variants of the exponential smoothing model depending on whether the researcher includes a trend in the model. We consider the following general form:[5]

$$f_t = f_{t-1} + T_{t-1} + \beta(y_{t-1} - f_{t-1}) \quad (5)$$

where: the trend function T_t can be: $T_t = 0$; $T_t = T_{t-1} + \gamma(y_{t-1} - f_{t-1})$; or $T_t = T_{t-1} + \gamma(y_{t-1} - f_{t-1})/f_{t-1}$. For each period, we estimate β and select the form of the trend from the variant providing the best in-sample fit.

3.2. New School Models

Andrews (1993) and Andrews and Ploeberger (1994) consider testing and estimating an unknown point change in a linear regression model. We consider the following linear regressions with a single break (2 regimes):

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \gamma_0 D_t + \varepsilon_t \quad (6a)$$

or:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + (\gamma_0 + \sum_{i=1}^{p} \gamma_i y_{t-i}) D_t + \varepsilon_t \quad (6b)$$

where: D_t is the Heaviside indicator indicating whether a break has occurred prior to period t.

In both cases we allow for lagged dependent variables as regressors to correct possible serial correlation. The first case (a) is a *partial* break model where the breaks are assumed to be in the intercept of the regression. The second case (b) is the case of a *pure* break model, where all

[5] In general, we use the symbol f_t^n to denote the n-step ahead forecast for period t. However, when it is unambiguous, or the number of steps ahead is unimportant, we drop the superscript n.

parameters are allowed to change across sample periods. In practice, it is necessary to 'trim' the data so that the various regimes in the breaking series have a sufficient number of observations to be properly estimated. We follow the conventional practice and use a trimming value $\varepsilon = 0.15$ so that each regime contains at least 15% of the observations. The null hypothesis of structural stability is tested against the alternative hypothesis of a one-time structural break using the supremum F (supF) tests of Andrews (1993).

If a break is detected in the pure break model, we re-estimate the series using only the post-break data. If a break is detected in the partial model, we forecast using two different types of equations. For the first type, we create a dummy variable and use the entire data set to estimate the model allowing the break to affect only the intercept term. For the second type, we re-estimate the model using only the post-break data.[6] Note that forecasts from this second method can differ from those of the pure break model since the AP test using the partial break model may not have the same size and power as the test using the pure break model.

Bai and Perron (1998) generalize the AP test by allowing for multiple structural breaks. We consider the following multiple linear regressions with m breaks (m + 1 regimes):

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + (\gamma_1 D_{1t} + \gamma_2 D_{2t} + ... + \gamma_m D_{mt}) + \varepsilon_t \tag{7a}$$

$$y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{m} D_{jt} (\gamma_{0j} + \sum_{i=1}^{p} \gamma_{ij} y_{t-i}) + \varepsilon_t \tag{7b}$$

$$y_t = \alpha_0 + \gamma_1 D_{1t} + \gamma_2 D_{2t} + ... + \gamma_m D_{mt} + u_t \tag{7c}$$

The first two cases allow for lagged dependent variables as regressors. The first case (a) is a *partial* break model where the breaks are assumed to be only in the intercept of the

[6] If we do not find a break we estimate and forecast using a simple *AR* (p) model.

regression. The second case (b) is the case of a *pure* break change model, where all parameters are allowed to change. The third case (c) is the case of a *pure* structural change mode accounting for possible serial correlation in a non-parametric way.[7] All models allow heterogeneity in the regression errors. Following Bai and Perron's (2003) recommendation to achieve tests with correct size in finite samples, we use a value of the trimming parameter $\varepsilon = 0.15$ and a maximum number of breaks $m = 5$.

The determination the appropriate number of breaks depends on the values of various tests statistics when the break dates are estimated. Specifically, we use the global and the sequential methods to estimate the number of breaks. First, we consider a supF type test of no structural change ($m = 0$) versus $m = k$ breaks, obtained by minimizing the global sum of squared residuals. To select the number of breaks, we follow the standard procedure and use the BIC. Second, we consider the sequential test, of ℓ versus $\ell+1$ breaks, labeled $F_t(\ell+1|\ell)$. For this test the first ℓ breaks are estimated and taken as given. The statistic $\sup F_t(\ell+1|\ell)$ is then calculated as the maximum of the *F*-statistics for testing no further structural change against the alternative of one additional change in the mean when the break date is varied over all possible dates. The procedure for estimating the number of breaks suggested by BP is based on the sequential application of the $\sup F_t(\ell+1|\ell)$ test. The procedure can be summarized as follows: begin with a test of no-break versus a single break. If the null hypothesis of no breaks is rejected, proceed to test the null of a single break versus two breaks, and so forth. This process is repeated until the statistics fail to reject the null hypothesis of no additional breaks. The estimated number of breaks is equal to the number of rejections.

[7] In this case, an estimate of the covariance matrix robust to heteroskedasticity and autocorrelation can be constructed using a quadratic spectral kernel and an AR (1) approximation to select the optimal bandwidth.

Since we consider models a, b and c using two testing procedures, we have six variants of the BP procedure. For example, the partial-sequential model uses 7a such that breaks are estimated sequentially, the pure-BIC model uses 7b such that breaks are estimated globally, and the *non-parametric* sequential model uses 7c such that breaks are estimated sequentially.

After estimating the timing of the breaks in each case we re-estimate the 6 models using the appropriate dummy variables indicated by the estimated break dates. As in the AP test, if we find breaks using the partial model, we estimate a model using the full data including the intercept dummies and a model using only the post-break data.

3.3 Nonlinear models: TAR and MTAR

In a sense, the New School models treat all breaks as permanent; a break can be ‘reversed’ only by a subsequent break of equal magnitude in the opposite direction. Moreover, even though multiple breaks occur, the mechanism generating the breaks is not estimated as part of the data-generating process. In contrast, regime switching models can be thought of as multiple-break models in which the breaking process is estimated along with the other parameters of the model. Although there are many types of regime switching models, we consider the threshold autoregressive model (TAR). The nature of the TAR model is that it allows for a number of different regimes with a separate autoregressive model in each regime. We will focus on the simple two-regime TAR model:

$$y_t = I_t\left[\alpha_{10} + \sum_{i=1}^{p}\alpha_{1i}y_{t-i}\right] + (1-I_t)\left[\alpha_{20} + \sum_{i=1}^{p}\alpha_{2i}y_{t-i}\right] + \varepsilon_t \tag{8}$$

$$I_t = \begin{cases} 1 & if \;\; y_{t-d} \geq \tau \\ 0 & if \;\; y_{t-d} < \tau \end{cases} \tag{9}$$

where: τ is the value of the threshold, p is the order of the model, d is the delay parameter, and I_t is the Heaviside indicator function.[8]

The nature of the TAR model is that there are two states of the world that we call 'recession' and 'expansion'. In a recession, the unemployment rate, y_{t-d}, exceeds the value of the threshold τ, so that I_t =1 and y_t follows the autoregressive process $\alpha_{10} + \sum \alpha_{1i} y_{t-i}$. Similarly, in an expansion, $y_{t-d} < \tau$, so that I_t = 0 and y_t follows the autoregressive process $\alpha_{20} + \sum \alpha_{2i} y_{t-i}$. Although y_t is linear in each regime, the possibility of regime switching means that the entire sequence is nonlinear.

The momentum threshold autoregressive (M-TAR) model used by Enders and Granger (1998) allows the regime to change according to the first difference of y_{t-d}. Hence, equation 9 is replaced with:

$$I_t = \begin{cases} 1 \;\; if \;\; \Delta y_{t-d} \geq \tau \\ 0 \;\; if \;\; \Delta y_{t-d} < \tau \end{cases} \qquad (10)$$

It is argued that the M-TAR model is useful for capturing situations in which the degree of autoregressive decay depends on the direction of change in y_t. Note that the M-TAR model should work well if the unemployment rate increases more readily than it decreases since persistence will depend on the value of Δy_{t-d}. Also note that if all $\alpha_{1i} = \alpha_{2i}$, the TAR and M-TAR models are equivalent to an $AR(p)$ model.

Both the TAR and M-TAR models permit us to estimate the value of the threshold without imposing an *a priori* line of demarcation between the regimes. The key feature of these models is that a sufficiently large shock can cause the system to switch between regimes. The

[8] We select p similar using the BIC.

dates at which the series crosses the threshold are not specified beforehand by the researcher; instead, the data determine whether the unemployment series is recession or expansion state.[9]

As analyzed in Koop, Pesaran, and Potter (1996), the multi-step ahead forecasts from a nonlinear model are state-dependent. Nevertheless, our out-of-sample forecasts with the TAR and M-TAR models can be constructed in such a way as to allow for the possibility of a regime change during the forecast horizon. As described in Enders (2004), we obtain the forecasts from the TAR and M-TAR models by simulation. Specifically**,** we select 12 randomly drawn realizations of the residuals of Equation 7. Because the residuals may not have a normal distribution, the residuals are selected using standard ‘‘bootstrapping’’ procedures. In particular, the residuals are drawn with replacement using a uniform distribution. We call these residuals $\varepsilon^*_{t+1}, \varepsilon^*_{t+2}, ..., \varepsilon^*_{t+12}$. We then generate y^*_{t+1} through y^*_{t+12} by substituting these ‘‘bootstrapped’’ residuals into Equation 7 and setting I_t appropriately for recessions or expansion states. For this particular history, we repeat the process 1000 times. The sample averages of y^*_{t+1} through y^*_{t+12} yield the one-step through 12-step ahead conditional forecasts of the unemployment series.

4. Comparative Performance of the Models

In this section we begin by consider expanding-window regressions to obtain multi-step-ahead forecasts from each of the estimated models. In our expanding-window regressions, we estimated the parameters of each model using all observations from the start of the series though 1986:4. We repeated the process by adding successive observations through 2004:6. The starting date of 1986:5 gives us a long span of data for each country and the ending date of 2004:6 allows

[9] A grid search over all potential values of the thresholds yields a superconsistent estimate of the unknown threshold parameter τ. We follow the conventional practice of excluding the highest and lowest 15% of the potential values to ensure an adequate number of observations on each side of the threshold. Note that our TAR amd M-TAR models constrain the variance of ε_t to be identical across the regimes.

us to compute 12-step ahead forecast errors through 2005:5.[10] At the end of this exercise there are 280 out-of-sample 1-step ahead through 12-step ahead forecasts for each country. The forecasts are used to obtain the mean (Bias), the mean absolute error (MAE), and the mean square prediction error (MSPE) of each method for each series at each forecasting horizon. Table 2 summarizes the models that have the smallest mean forecast error (bias), mean absolute errors (MAE) and mean square prediction errors (MSPE) against the linear model at 1, 3, 6 and 12 steps ahead for each country (except the nonlinear model). A full reporting of the performance of the various models is contained in the Appendix.

4.1. Forecast performance of the models with an expanding window

In Table 2, and in the Appendix, we report the forecasting performance of the following models (described in detail in section 3) using an expanding window. Our key findings are as follows:

1. If the objective is to minimize the MAE or the MSPE, the linear model generally works very well at any forecast horizon. In fact, for the United States, the linear model has the smallest MAE and MSPE at every forecast horizon. For the other countries, the linear model is usually within a few percentage points of the best models.

2. If the objective is to minimize the bias, second differencing works in almost all cases. This should not be too surprising since, as argued by Clements and Hendry (1999), second differencing acts to eliminate the effects of structural breaks and removes persistence form the series. For Canada, Australia, Denmark, and France, exponential smoothing produces the smallest bias at all forecasting horizons. For the US, Japan and Germany, it produce the smallest bias at 1-step, 3-step, and 6-step ahead forecasting horizons. On the other hand, second differencing very rarely minimizes the MSPE and MAE.

[10] We used several forecasting origins past 1986:4 but the results are similar to those discussed here.

3. At very short forecasting horizons, exponential smoothing generally works quite well and sometimes results in the best values of the MSPE and MAE. Note that exponential smoothing results in the smallest bias, MAE and MSPE in almost all cases for Netherlands and Norway. However, in general, it does not work well at long horizons for the measures of dispersion (MAE and MSPE).

4. In contrast to the argument of Diebold and Kilian (2000), pre-testing does not seem to work better than simply estimating a model in levels or in first-differences. In most cases, the bias, MAE or the MSPE from pre-testing are equal to or greater than those of the linear model or the first-differenced model. A possible explanation is low power of unit root tests when the data includes structural change; it is possible to incorrectly find a unit root when data is in fact regime-wise stationary.

5. A key finding is that tests for structural change rarely outperform the Old School models for bias, MAE, or MSPE at any horizon. At the first- and third- step ahead they almost never deliver good out-of-sample forecasts. There are a few instances in which they have the best performance (France, Japan and UK), but there is no clear pattern as to which variant is best. In comparing the one-break AP test to the BP tests, it seems that the multiple structural change tests perform slightly better.

6. Although OS methods deem to dominate NS methods, it is of interest to determine which of the various New School methods works the best. As detailed in the Appendix, the parametric method seem to give a better performance than the nonparametric method. Also note that methods using the BIC generally work better than the sequential method. There is no clear pattern as to whether the partial models work better then the pure break models.

There are several possible reasons why the various NS methods perform poorly. First, tests for structural change tests deliver a poor in-sample performance when analyzing highly persistent data, due to the large size distortions of the test [see Diebold and Chen (1996) and Prodan (2005)]. In fact, many of the NS models do best when they detect no breaks. Second, the sequential method lacks power if there are breaks of opposite sign. Since the unemployment rates do not appear to trend on one direction or the other, any multiple breaks in the rates are likely to be of opposite sign. Third, Elliott (2005) presents analytical results that forecasts based on point estimates of the break dates are unlikely to improve forecasts. He argues that if breaks are small relative to the variance of the series, least square methods will not provide a consistent estimate of the break point. Simply ignoring the break can provide better forecasts that an incorrect selection of the break date. Hence, it is not surprising that, in such circumstances, OS methods of correcting for breaks can do better than the NS methods.

4.2. New School methods using only post break methods

In the Appendix, we report the effects of using only the post-break data. Our conclusion is that there are not any important patterns in the forecasting performance of the New School models estimated over the full sample period versus those using only the post-break data. As such, it cannot be claimed that one of these NS methods is preferable to the other. For example, in the case of the U.S. using the sequential method and the partial break model, the MSPEs for 1-, 3-, 6- and 12-step ahead forecasts are 1.0678, 1.2537,1.5348, and 1.8477 times those of the linear model, respectively. As shown in the Appendix, these numbers fall to 1.0433, 1.8565, 1.3922 and 1.6599 if only the post-break data is used. However, using only the post-break data increases the MSPE at all forecasting horizons if the sequential method in the pure break model is used. The opposite pattern holds for Canada.

4.3. Forecasting performance of the TAR and M-TAR Models

In Table 3 we compare the forecasting performance of the nonlinear models (TAR and MTAR) against the linear model. We find that TAR model does not improve the out of sample forecasting performance over the linear model, except in few cases when improving the bias (for cases as France, Germany, Japan and US). Previously, Rothman (2001) argued that improved forecasting performance can be achieved by using nonlinear time series models for US unemployment rate, finding support for the claim that nonlinear forecast can improve linear forecasts. For the untransformed series, a simple linear model usually performed the better than the threshold models.

4.4. Forecasting Performance with Intercept Corrections

For each of the OS forecasting methods discussed above, we constructed the intercept corrections using the simple average of the previous twelve forecast errors. To explain, let e_t^n denote the *n*-step ahead forecast error for period *t* constructed as $y_t - f_t^n$. The intercept correction for the *n*-step ahead forecast of y_t (IC_t) can be constructed as:

$$IC_t^n = \sum_{t-12}^{t-1} e_t^n / 12 \qquad (11)$$

The value of the intercept correction (IC) will be positive (negative) if the model has, on average, under-forecasted (over-forecasted) during the previous twelve months. In a sense, IC_t^n is the bias in the *n*-step ahead forecasts over the previous year. To use the intercept correction, simply add IC_t^n to f_t^n .

Perhaps the key lesson of the paper is that end point corrections work extremely well for long-term forecasting. Specifically, the MAE and the MSPE of the intercept corrected forecasts using the linear model at twelve month horizons are always lower than those of any other model! Moreover, except for the UK, the MAE and the MSPE of the intercept corrected forecasts using the linear model at six month horizons are always lower than those of any other model! In many instances, forecasts from an estimated model have more than twice the MSPE of the forecasts from the same model using the IC. For example, the MSPE of the 12-step ahead forecasts of the IC model relative to that of the simple linear model is 49.4% for the US, 20.3% for the Netherlands, 40.1% for Japan, and 38.2% for Norway. Although not reported in the Appendix, intercept corrections often performed well in NS models. In a sense, an IC acts as a coefficient change occurring near the end of the estimation period.

The reason that the intercept correction works so well is that the forecast errors for each unemployment series are strongly serially correlated at the longer-term forecasts. For example, the first six autocorrelations (ρ_i) of the US forecast errors from the $AR(p)$ model are:

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
1-Step	-0.126	-0.183	-0.019	0.062	0.089	0.031
3-Step	0.474	0.130	-0.151	0.073	0.213	0.184
6-Step	0.729	0.567	0.489	0.430	0.344	0.190
12-Step	0.907	0.848	0.799	0.759	0.719	0.662

Given that the correlation coefficient between the 12-step ahead forecast errors for periods t and t-1 is 0.907, if the recent 12-step forecasts have been too low (too high), the subsequent forecasts should be adjusted upward (downward).[11] In contrast, the correlations between the 1-step ahead forecast are low so that the intercept correction does not improve on the forecasting performance of the model.

The lessons about the intercept correction can be summarized as:

1. Always use the an intercept correction for long-term forecasting using a $AR(p)$ model or an ARI(p, 1) model.

2. Do not use an intercept correction with exponential smoothing; in most instances, the IC acts to increase the MSPE. The very nature of exponential smoothing is designed to take the most recent forecast errors into consideration when making subsequent forecasts.

3. In general**,** do not combine intercept corrections with second differencing. An intercept in a model with second differences represents a quadratic time trend in the level of the series. Any adjustments to a term multiplying a squared time trend are likely to induce large long-term out-of-sample forecast errors.

4.5. Encompassing: Combining Old School and New School Models

It is well-known that combinations of forecasts formed from alternative models often yield far better forecasts than any one individual forecast. That is certainly true of the unemployment models too. Since our focus is on Old School versus New School forecasts, we will concentrate on the issue of whether forecasts using the Bai-Perron methodology encompass

[11] We report results using a 12-month window to construct the intercept correction since we use monthly data. We experimented with shorter windows and the results were quite similar to those reported here. Moreover, in (11), each forecast error is given a weight of 1/12. It is possible to construct intercept corrections that give higher weights to the more recent forecast errors.

those from the linear model. One might expect the forecasts from models with estimated breaks to incorporate all of the information contained in a simple linear model that ignores breaks. If this way of thinking is correct, forecasts from the BP models should encompass those of the linear model.

Chong and Hendry (1986) propose a simple encompassing test. Although a number of similar tests have been proposed in the recent literature, the Chong-Hendry test is simple to understand, convenient to use with a large number of series, and gives us a sense of the magnitude of the encompassing, if any. Let f_{Bt} denote the forecast from a model estimated using the BP procedure and let f_t denote the forecast from the linear model. The forecasts of the two models can be combined to form the pooled forecast (f_{Pt}) as:

$$f_{Pt} = (1 - \gamma)f_{Bt} + \gamma f_t \tag{12}$$

or

$$y_t = (1 - \gamma)f_{Bt} + \gamma f_t + e_{Pt} \tag{13}$$

where: γ is a weight and e_{Pt} is the forecast error of the pooled forecast constructed as $y_t - f_{Pt}$.

If the linear model adds no information to the forecast of the BP model, $\gamma = 0$. Hence, a straightforward test of encompassing can be obtained by forming the forecast error of the BP model as $e_{Bt} = y_t - f_{Bt}$ so that (13) can be rewritten as:

$$e_{Bt} = \gamma(f_t - f_{Bt}) + e_{Pt} \tag{14}$$

The test consists of regressing the differential forecasts $f_t - f_{Bt}$ on the forecast errors from the BP model. If a t-test for the null hypothesis $\gamma = 0$ can be rejected, it can be concluded that information from the linear model acts to improve the forecasts of the New School models. Of course, for some models, the test is meaningless since the BP model indicated no breaks. Nevertheless, for other cases, Table 4 indicates that almost every t-statistic (calculated using robust standard errors) at every forecasting horizon for every variant of the BP model was of sufficient magnitude to reject the null hypothesis $\gamma = 0$.[12] As such, the forecasts from the linear model add meaningful information to forecasts from the BP model. It must be concluded that the BP procedure (in all of its variants) either yields too many breakpoints, yields poor estimates of the break dates, and/or yields poor estimates of the magnitudes of the breaks.

5. Conclusion

This study presents an in-depth examination of the forecasts for the monthly unemployment rates in 10 OECD countries, using various time series models. Ignoring the breaks and using a simple linear model often does quite well if the goal is to minimize the MSPE or the MAE. The linear model is the best for United States (at all steps) and Canada (except for the first step, where the exponential smoothing outperforms the linear model) and works very well for Australia. An examination of the graphs of these series (see Figure 2), suggests that the common feature for these countries is that their unemployment rates seem to fluctuate around a constant mean, possible being stationary. If there are breaks in these series, they are probably small and possibly offsetting. The *first difference* forecasts generated low values of the MAE and MSPE for Germany at all forecasting horizons. Also, at 6 and 12 steps, first differencing performed very well for countries as Denmark and Australia. The unemployment rates series for Denmark and Germany, have a U-shape and seem to have a unit root in their data generating process.

[12] We report results using the full data set. Results using the post break data are quite similar. Note that it is possible to use a full break model to detect the break and to forecast with a model that allows only breaks in the intercept. The results of such forecasts are contained in columns 5 and 6 of the table and are dubbed AllSeq(c) and AllBIC(c).

Exponential smoothing works very well at short horizons for several countries. Note that exponential smoothing generates the lowest MAE and MSPE, at all steps, for Netherlands and Norway. In both cases, the unemployment rates seem to include a time-varying trend. For countries as France and Japan, either *second differencing* or the *structural change tests* works well when out-of-sample forecasting. In both cases, Japan and France, second differencing is outperformed by multiple structural change tests at the 12th step.

In the case of UK there is no specific pattern: as in several other cases exponential smoothing works well at the first step ahead, followed by structural change tests and first differencing at 3rd, 6th and 12th step ahead. We are able to draw some recommendations when estimating persistent data with possible breaks[13]:

1. If the goal is to minimize dispersion (MAE, MSPE), use exponential smoothing when forecasting at short horizons.

2. If the goal is to minimize dispersion (MAE, MSPE), use the linear model with an intercept correction for forecasting at longer horizons. .

3. Use second differencing at any horizon to minimize the bias.

4. For this data set, OS models generally perform better than NS models. However, if one chooses to use structural change tests, use the BP test selecting the breaks using BIC and correct for serial correlation using the parametric method.

5. Even though OS methods seem to win the forecasting competition, there is no reason to forecast using only one method. We presented evidence that combining OS and NS methods can improve forecast accuracy.

[13] We are aware that our recommendations are based on analyzing specific data. A Monte Carlo simulation exercise is needed in order to assess the robustness of our results and this is subject of our future research.

Appendix

For each country and forecasting technique, the tables below report the bias, mean absolute error (MAE), and mean square prediction error (MSPE) at the forecast horizons of 1 step ahead, 3 steps ahead, 6 steps ahead, and 12 steps ahead.

In order to condense the tables, the forecasts of the various models are denoted by the following terms:

Linear:	the $AR(p)$ model
Diff:	the $AR(p, 1)$ model
Pretest:	forecasts based on pretesting for a unit root
2nd Diff:	the $AR(p, 2)$ model
Smooth:	exponential smoothing
AP:	pure break form of the AP test
AP(c):	the partial break form of the AP test
PSeq:	the partial break form of BP test using the sequential method
PBIC:	the partial break form of BP test using the global method
Seq:	the pure break form of BP test using the sequential method
BIC:	the pure break form of BP test using the global method
NPSeq:	the nonparametric form of the BP test using the sequential method
NPBIC:	the nonparametric form of the BP test using the global method

The linear model is the benchmark model so each entry is relative to that of the linear forecast. For example, in the case of the United States, the bias of the of the 1 step ahead forecasts using first differences is 0.6543 times that of the linear model. Similarly, the entries for AP, PBIC and BIC are all 1.0000 since these methods found no breaks for the U.S.

In the top portion of the tables, all forecasts were obtained using the full data set. Boldface entries indicate the lowest entry in a row. In this portion of the table, a row not boldfaced indicates that the linear model performed best. The lower left portion contains the forecasts using intercept corrections for the linear, Diff, Pretest, 2nd Diff and Smooth. Note that that the entries are all relative to the linear model without an intercept correction. Hence, the IC reduced the MSPE 12-step ahead forecasts of the linear model to 0.494 of the original value. The lower right portion of the table contains the forecasts (relative to the linear model) for the NS methods using only the post-break data. Since, except for the treatment of lags, forecasts from the pure break model are invariant to the choice of pre-break and post-break data, the post-break values for AP, Seq and BIC using only post-break data are not listed in the table.

United States

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
							Full Data Set						
1 Step Ahead													
Mean		0.654	1.227	**0.129**	0.168	1.000	3.426	3.994	1.000	0.965	1.000	3.822	3.354
MAE		1.006	1.001	1.027	1.016	1.000	1.028	1.029	1.000	1.002	1.000	1.028	1.049
MSPE		1.011	1.007	1.075	1.037	1.000	1.068	1.068	1.000	1.004	1.000	1.072	1.122
3 Step Ahead													
Mean		0.647	1.222	**0.071**	0.106	1.000	3.458	3.843	1.000	0.967	1.000	3.675	3.229
MAE		1.021	1.013	1.099	1.069	1.000	1.124	1.146	1.000	1.004	1.000	1.158	1.234
MSPE		1.038	1.027	1.226	1.153	1.000	1.231	1.254	1.000	1.013	1.000	1.282	1.490
6 Step Ahead													
Mean		0.640	1.209	**0.043**	0.073	1.000	3.411	3.848	1.000	0.967	1.000	3.675	3.216
MAE		1.004	1.032	1.182	1.125	1.000	1.262	1.279	1.000	1.006	1.000	1.333	1.499
MSPE		1.088	1.057	1.539	1.398	1.000	1.541	1.535	1.000	1.024	1.000	1.623	2.087
12 Step Ahead													
Mean		0.621	1.198	-0.030	**-0.002**	1.000	3.341	3.747	1.000	0.968	1.000	3.543	3.040
MAE		1.035	1.042	1.353	1.287	1.000	1.365	1.408	1.000	1.007	1.000	1.379	1.621
MSPE		1.185	1.086	2.114	1.879	1.000	1.819	1.848	1.000	1.023	1.000	1.873	2.534
			End Point Correction						**Post-Break Data**				
1 Step Ahead													
Mean	-0.092	-0.192	-0.138	-0.142	-0.125		3.036	3.393	1.000			1.667	-0.890
MAE	1.035	1.040	1.035	1.077	1.063		1.015	1.012	1.000			1.008	1.048
MSPE	1.057	1.072	1.058	1.189	1.155		1.051	1.043	1.000			1.049	1.148
3 Step Ahead													
Mean	-0.103	-0.204	-0.153	-0.161	-0.145		3.133	3.361	1.000			1.825	0.012
MAE	0.986	1.011	0.989	1.177	1.151		1.101	1.111	1.000			1.073	1.129
MSPE	0.977	1.028	0.980	1.424	1.342		1.183	1.187	1.000			1.156	1.346
6 Step Ahead													
Mean	-0.094	-0.192	-0.143	-0.144	-0.132		3.131	3.413	1.000			1.920	0.383
MAE	0.837	0.891	0.842	1.294	1.245		1.216	1.217	1.000			1.177	1.237
MSPE	0.793	0.903	0.801	1.845	1.675		1.429	1.392	1.000			1.381	1.617
12 Step Ahead													
Mean	-0.088	-0.186	-0.138	-0.131	-0.121		3.129	3.407	1.000			1.972	0.678
MAE	0.669	0.778	0.680	1.444	1.340		1.323	1.340	1.000			1.224	1.325
MSPE	0.494	0.656	0.507	2.407	2.058		1.684	1.660	1.000			1.602	1.819

CANADA

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
							Full Data Set						
1 Step Ahead													
Mean		2.617	2.617	-0.179	4.892	3.471	3.067	1.459	1.000	-2.462	1.000	2.963	1.960
MAE		0.987	0.987	1.037	0.986	1.046	1.022	1.001	1.000	1.023	1.000	1.018	1.072
MSPE		0.999	0.999	1.064	0.994	1.059	1.028	1.007	1.000	1.017	1.000	1.020	1.098
3 Step Ahead													
Mean		2.847	2.847	-0.292	5.403	3.769	3.325	1.453	1.000	-2.341	1.000	2.941	1.945
MAE		0.999	0.999	1.103	1.039	1.109	1.067	1.013	1.000	1.040	1.000	1.069	1.185
MSPE		1.008	1.008	1.248	1.086	1.196	1.082	1.025	1.000	1.079	1.000	1.063	1.315
6 Step Ahead													
Mean		2.751	2.751	-0.253	5.055	3.899	3.427	1.560	1.000	-2.899	1.000	3.379	2.196
MAE		1.027	1.027	1.286	1.152	1.176	1.116	1.026	1.000	1.102	1.000	1.108	1.332
MSPE		1.021	1.021	1.565	1.255	1.448	1.159	1.042	1.000	1.187	1.000	1.133	1.634
12 Step Ahead													
Mean		2.644	2.644	-0.284	4.775	4.098	3.449	1.634	1.000	-3.299	1.000	3.556	2.369
MAE		1.038	1.038	1.520	1.232	1.259	1.172	1.052	1.000	1.183	1.000	1.162	1.451
MSPE		1.033	1.033	2.172	1.376	1.916	1.251	1.073	1.000	1.374	1.000	1.197	1.874
			End Point Correction				**Post-Break Data**						
1 Step Ahead													
Mean	-0.414	-0.293	-0.293	-0.690	-0.505		3.029	1.133	1.000			5.004	2.931
MAE	1.052	1.044	1.044	1.105	1.037		1.027	1.007	1.000			1.052	1.084
MSPE	1.087	1.083	1.083	1.205	1.097		1.022	1.015	1.000			1.061	1.124
3 Step Ahead													
Mean	-0.618	-0.480	-0.480	-1.032	-0.797		3.371	1.221	1.000			5.564	3.394
MAE	1.026	1.025	1.025	1.225	1.079		1.075	1.023	1.000			1.141	1.200
MSPE	1.043	1.041	1.041	1.497	1.206		1.097	1.045	1.000			1.252	1.397
6 Step Ahead													
Mean	-1.894	-1.918	-1.918	-2.686	-2.290		3.456	1.333	1.000			6.663	4.278
MAE	0.994	0.999	0.999	1.470	1.163		1.126	1.031	1.000			1.281	1.399
MSPE	0.923	0.925	0.925	2.018	1.386		1.183	1.055	1.000			1.487	1.841
12 Step Ahead													
Mean	3.908	4.319	4.319	5.156	4.607		3.333	1.453	1.000			6.940	4.374
MAE	0.862	0.862	0.862	1.795	1.062		1.164	1.053	1.000			1.349	1.549
MSPE	0.643	0.647	0.647	2.844	1.286		1.284	1.084	1.000			1.673	2.248

AUSTRALIA

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		1.272	1.272	-0.025	0.059	1.435	1.474	0.862	0.435	2.163	1.000	1.934	2.429
MAE		0.995	0.995	0.999	0.977	1.001	1.009	1.002	0.986	0.997	1.000	1.018	1.008
MSPE		1.017	1.017	1.038	0.990	1.009	1.014	1.001	0.963	0.986	1.000	1.066	1.050
3 Step Ahead													
Mean		1.132	1.132	-0.033	0.054	1.396	1.423	0.842	0.335	2.096	1.000	1.903	2.402
MAE		1.002	1.002	1.077	1.027	1.035	1.038	1.003	1.003	1.032	1.000	1.115	1.066
MSPE		1.013	1.013	1.137	1.051	1.045	1.052	1.003	1.002	1.016	1.000	1.244	1.143
6 Step Ahead													
Mean		1.028	1.028	0.084	0.146	1.316	1.322	0.866	0.345	1.991	1.000	1.665	2.257
MAE		0.964	0.964	1.090	1.037	1.047	1.048	0.995	1.026	1.061	1.000	1.205	1.155
MSPE		1.023	1.023	1.289	1.184	1.095	1.093	0.997	1.166	1.095	1.000	1.544	1.324
12 Step Ahead													
Mean		0.991	0.991	0.187	0.234	1.243	1.183	0.884	0.356	1.974	1.000	1.363	2.190
MAE		0.911	0.911	1.143	1.107	1.046	1.035	1.009	1.038	1.115	1.000	1.178	1.203
MSPE		0.985	0.985	1.542	1.438	1.128	1.059	1.016	1.251	1.209	1.000	1.535	1.467
			End Point Correction					**Post-Break Data**					
1 Step Ahead													
Mean	0.235	0.085	0.085	-0.048	-0.046		1.421	0.834	0.560			3.061	4.249
MAE	1.016	1.030	1.030	1.053	1.028		1.010	1.001	1.001			1.076	1.099
MSPE	1.085	1.124	1.124	1.154	1.109		1.013	1.000	0.999			1.252	1.330
3 Step Ahead													
Mean	0.086	-0.072	-0.072	-0.209	-0.215		1.379	0.854	0.526			2.333	3.363
MAE	0.998	1.030	1.030	1.137	1.085		1.034	1.003	1.008			1.205	1.257
MSPE	1.001	1.079	1.079	1.252	1.169		1.044	1.002	1.022			1.474	1.606
6 Step Ahead													
Mean	0.067	-0.093	-0.093	-0.231	-0.231		1.297	0.890	0.522			1.956	2.353
MAE	0.835	0.905	0.905	1.112	1.072		1.044	0.989	1.028			1.325	1.362
MSPE	0.760	0.945	0.945	1.315	1.229		1.083	0.993	1.129			1.773	1.993
12 Step Ahead													
Mean	0.107	-0.063	-0.063	-0.207	-0.207		1.176	0.918	0.492			1.577	1.571
MAE	0.633	0.753	0.753	1.116	1.056		1.034	0.998	1.040			1.367	1.412
MSPE	0.480	0.711	0.711	1.395	1.271		1.056	0.999	1.208			1.916	2.295

NETHERLANDS

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		0.565	0.565	-0.046	-0.013	0.877	0.575	0.662	1.000	0.972	1.573	-0.142	-0.215
MAE		1.003	1.003	1.011	0.980	0.998	1.017	1.004	1.000	0.998	1.103	1.018	1.015
MSPE		0.977	0.977	0.983	0.960	1.000	1.024	1.012	1.000	0.994	1.141	0.983	0.972
3 Step Ahead													
Mean		0.571	0.571	-0.025	0.007	0.868	0.583	0.662	1.000	0.976	1.522	-0.148	-0.218
MAE		1.001	1.001	0.945	0.912	0.960	1.061	1.037	1.000	1.001	1.203	0.987	0.958
MSPE		0.951	0.951	0.907	0.794	0.948	1.058	1.031	1.000	0.993	1.332	0.955	0.928
6 Step Ahead													
Mean		0.565	0.565	-0.018	0.008	0.914	0.582	0.657	1.000	0.980	1.464	-0.168	-0.235
MAE		0.968	0.968	0.865	0.804	0.930	1.082	1.059	1.000	1.002	1.248	0.963	0.911
MSPE		0.911	0.911	0.786	0.644	0.917	1.104	1.061	1.000	1.002	1.507	0.942	0.883
12 Step Ahead													
Mean		0.561	0.561	-0.035	-0.009	1.101	0.578	0.655	1.000	0.983	1.379	-0.190	-0.252
MAE		0.987	0.987	0.784	0.778	1.069	1.072	1.049	1.000	1.007	1.224	0.939	0.887
MSPE		0.929	0.929	0.769	0.669	1.214	1.148	1.099	1.000	1.007	1.497	1.010	0.924
			End Point Correction					Post-Break Data					
1 Step Ahead													
Mean	-0.074	-0.116	-0.116	-0.016	-0.022		0.563	0.602	1.000			-0.031	-0.361
MAE	0.990	0.990	0.990	1.056	1.005		1.020	1.015	1.000			1.060	1.050
MSPE	0.948	0.952	0.952	1.085	1.025		1.024	1.018	1.000			1.041	1.046
3 Step Ahead													
Mean	-0.051	-0.092	-0.092	0.015	0.004		0.571	0.604	1.000			0.048	-0.352
MAE	0.842	0.849	0.849	0.977	0.882		1.057	1.048	1.000			1.064	1.071
MSPE	0.720	0.736	0.736	0.996	0.791		1.053	1.046	1.000			1.072	1.078
6 Step Ahead													
Mean	-0.057	-0.098	-0.098	0.015	0.002		0.574	0.609	1.000			0.129	-0.344
MAE	0.587	0.603	0.603	0.875	0.688		1.085	1.080	1.000			1.071	1.048
MSPE	0.380	0.399	0.399	0.837	0.521		1.097	1.086	1.000			1.161	1.092
12 Step Ahead													
Mean	-0.067	-0.109	-0.109	0.013	-0.002		0.571	0.611	1.000			0.191	-0.331
MAE	0.410	0.422	0.422	0.786	0.584		1.068	1.058	1.000			1.081	0.999
MSPE	0.203	0.220	0.220	0.772	0.399		1.138	1.123	1.000			1.270	1.067

DENMARK

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		0.548	0.548	0.017	-0.365	0.221	0.263	-1.138	-0.393	-1.390	0.597	1.049	0.474
MAE		0.983	0.983	0.990	0.964	1.026	1.024	1.029	1.048	0.993	1.015	1.025	1.049
MSPE		0.992	0.992	1.061	1.002	1.037	1.036	1.057	1.083	1.002	1.012	1.059	1.103
3 Step Ahead													
Mean		0.575	0.575	0.065	-0.244	0.345	0.378	-0.880	-0.107	-1.163	0.643	1.043	0.528
MAE		0.975	0.975	1.060	0.987	1.060	1.055	1.124	1.122	1.022	1.025	1.090	1.159
MSPE		0.980	0.980	1.144	0.980	1.081	1.076	1.132	1.171	1.018	1.032	1.152	1.287
6 Step Ahead													
Mean		0.568	0.568	0.028	-0.227	0.463	0.424	-0.794	0.061	-1.112	0.653	1.047	0.541
MAE		0.962	0.962	1.070	0.970	1.082	1.091	1.200	1.178	1.053	1.040	1.155	1.312
MSPE		0.973	0.973	1.199	0.976	1.119	1.135	1.235	1.249	1.042	1.058	1.282	1.555
12 Step Ahead													
Mean		0.605	0.605	0.082	-0.112	0.930	0.555	-0.518	0.364	-0.879	0.694	1.045	0.614
MAE		0.943	0.943	1.024	0.927	1.027	1.116	1.231	1.179	1.056	1.063	1.206	1.377
MSPE		0.985	0.985	1.230	0.997	1.035	1.208	1.350	1.297	1.093	1.099	1.455	1.824
			End Point Correction					**Post-Break Data**					
1 Step Ahead													
Mean	0.214	0.179	0.179	0.296	0.398		0.369	-1.402	-1.384			-0.564	-1.214
MAE	0.988	0.994	0.994	1.066	1.026		1.020	1.004	1.131			1.065	1.061
MSPE	1.028	1.038	1.038	1.179	1.089		1.029	1.043	1.265			1.196	1.166
3 Step Ahead													
Mean	0.178	0.144	0.144	0.237	0.348		0.461	-0.867	0.043			-0.055	-0.593
MAE	0.946	0.963	0.963	1.136	1.014		1.045	1.124	1.176			1.196	1.154
MSPE	0.914	0.937	0.937	1.281	1.048		1.063	1.133	1.359			1.473	1.434
6 Step Ahead													
Mean	0.139	0.105	0.105	0.186	0.289		0.495	-0.549	0.723			0.184	-0.255
MAE	0.821	0.851	0.851	1.133	0.991		1.077	1.181	1.196			1.253	1.225
MSPE	0.697	0.744	0.744	1.312	0.978		1.115	1.213	1.446			1.644	1.624
12 Step Ahead													
Mean	0.154	0.123	0.123	0.193	0.274		0.598	-0.174	1.157			0.440	0.169
MAE	0.579	0.638	0.638	1.034	0.850		1.104	1.188	1.219			1.286	1.241
MSPE	0.387	0.460	0.460	1.231	0.813		1.184	1.305	1.544			1.754	1.728

FRANCE

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		1.386	1.386	0.099	0.426	1.000	1.000	0.965	0.865	0.483	1.000	0.836	0.548
MAE		1.008	1.008	0.988	0.997	1.000	1.000	1.018	0.995	1.009	1.000	0.994	1.009
MSPE		1.007	1.007	0.994	0.928	1.000	1.000	1.026	0.991	1.028	1.000	1.020	1.026
3 Step Ahead													
Mean		1.399	1.399	0.061	0.431	1.000	1.000	0.927	0.851	0.442	1.000	0.844	0.539
MAE		1.011	1.011	0.947	1.026	1.000	1.000	1.039	0.998	1.039	1.000	1.013	1.048
MSPE		1.018	1.018	1.011	1.121	1.000	1.000	1.054	0.986	1.058	1.000	1.045	1.068
6 Step Ahead													
Mean		1.398	1.398	0.071	0.354	1.000	1.000	0.897	0.841	0.427	1.000	0.855	0.547
MAE		1.021	1.021	0.910	0.977	1.000	1.000	1.060	0.991	1.042	1.000	1.014	1.062
MSPE		1.032	1.032	1.040	1.157	1.000	1.000	1.089	0.979	1.091	1.000	1.057	1.103
12 Step Ahead													
Mean		1.407	1.407	0.065	0.289	1.000	1.000	0.847	0.814	0.390	1.000	0.865	0.554
MAE		1.033	1.033	0.974	1.028	1.000	1.000	1.063	0.957	1.050	1.000	1.003	1.066
MSPE		1.056	1.056	1.207	1.312	1.000	1.000	1.117	0.950	1.120	1.000	1.066	1.155
			End Point Correction					**Post-Break Data**					
1 Step Ahead													
Mean	-0.078	-0.070	-0.070	-0.017	-0.101		1.000	0.819	0.849			1.085	0.500
MAE	1.007	1.008	1.008	1.062	1.052		1.000	1.005	0.997			1.075	1.090
MSPE	0.996	0.996	0.996	1.078	1.014		1.000	1.008	0.995			1.078	1.126
3 Step Ahead													
Mean	-0.163	-0.156	-0.156	-0.099	-0.174		1.000	0.731	0.834			1.360	0.746
MAE	0.905	0.903	0.903	1.024	1.087		1.000	1.040	0.994			1.212	1.270
MSPE	0.885	0.884	0.884	1.107	1.234		1.000	1.067	0.981			1.391	1.468
6 Step Ahead													
Mean	-0.165	-0.158	-0.158	-0.100	-0.163		1.000	0.701	0.832			1.318	0.753
MAE	0.796	0.801	0.801	0.972	1.031		1.000	1.071	0.986			1.256	1.306
MSPE	0.759	0.757	0.757	1.138	1.247		1.000	1.138	0.972			1.615	1.710
12 Step Ahead													
Mean	-0.232	-0.226	-0.226	-0.162	-0.216		1.000	0.633	0.812			1.137	0.587
MAE	0.676	0.681	0.681	0.986	1.009		1.000	1.099	0.956			1.239	1.321
MSPE	0.579	0.576	0.576	1.260	1.310		1.000	1.240	0.946			1.845	2.006

JAPAN

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		2.032	2.032	-0.630	0.477	1.276	1.116	1.032	0.176	1.000	1.000	1.652	2.313
MAE		0.997	0.997	0.988	0.998	0.987	0.998	0.995	1.035	1.000	1.000	1.007	1.117
MSPE		0.997	0.997	0.942	0.993	0.970	0.995	0.985	1.044	1.000	1.000	1.001	1.156
3 Step Ahead													
Mean		2.107	2.107	-0.611	0.466	1.329	1.152	1.056	0.262	1.000	1.000	1.637	2.285
MAE		0.997	0.997	0.919	0.967	0.965	0.991	0.982	1.055	1.000	1.000	1.003	1.202
MSPE		0.991	0.991	0.843	0.951	0.928	0.982	0.958	1.080	1.000	1.000	0.993	1.327
6 Step Ahead													
Mean		2.154	2.154	-0.555	0.453	1.411	1.186	1.094	0.249	1.000	1.000	1.712	2.455
MAE		0.979	0.979	0.876	0.965	0.950	0.981	0.971	1.094	1.000	1.000	1.019	1.284
MSPE		0.968	0.968	0.778	0.963	0.902	0.968	0.938	1.110	1.000	1.000	0.994	1.437
12 Step Ahead													
Mean		2.515	2.515	-0.892	0.309	1.686	1.310	1.205	0.148	1.000	1.000	1.963	2.993
MAE		0.956	0.956	0.991	0.969	0.972	0.973	0.964	1.076	1.000	1.000	1.014	1.230
MSPE		0.916	0.916	0.972	0.980	0.951	0.943	0.919	1.080	1.000	1.000	0.985	1.328
			End Point Correction					**Post-Break Data**					
1 Step Ahead													
Mean	-1.683	-1.535	-1.535	-0.719	-1.530		0.715	0.306	0.147			0.326	-1.182
MAE	0.999	0.993	0.993	1.054	1.010		0.998	1.000	1.027			1.007	1.061
MSPE	0.960	0.954	0.954	1.029	0.969		1.000	0.987	1.019			1.000	1.056
3 Step Ahead													
Mean	-0.837	-0.727	-0.727	-0.069	-0.737		0.667	0.284	0.121			0.269	-1.193
MAE	0.870	0.855	0.855	0.947	0.853		0.997	0.969	1.029			0.968	1.042
MSPE	0.766	0.749	0.749	0.900	0.733		1.008	0.959	1.032			0.953	1.051
6 Step Ahead													
Mean	-0.495	-0.398	-0.398	0.201	-0.409		0.655	0.193	0.105			0.168	-1.356
MAE	0.753	0.729	0.729	0.914	0.753		1.015	0.981	1.050			0.979	1.057
MSPE	0.559	0.524	0.524	0.792	0.555		1.030	0.964	1.048			0.948	1.060
12 Step Ahead													
Mean	-0.310	-0.213	-0.213	0.409	-0.226		0.580	-0.069	-0.042			-0.104	-2.007
MAE	0.616	0.576	0.576	0.904	0.627		1.035	1.026	1.045			1.021	1.092
MSPE	0.401	0.345	0.345	0.870	0.416		1.037	1.004	1.030			0.977	1.051

NORWAY

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
							Full Data Set						
1 Step Ahead													
Mean		0.056	0.139	-0.288	-0.016	-1.483	1.929	1.157	0.119	0.487	0.221	0.049	-0.733
MAE		0.995	0.995	1.011	0.950	0.981	1.036	1.014	0.997	1.021	1.032	1.050	1.003
MSPE		0.990	0.988	1.055	0.945	0.979	1.018	1.021	1.022	1.041	1.068	1.083	1.021
3 Step Ahead													
Mean		-0.214	-0.120	-0.295	0.072	-1.789	2.055	1.204	-0.041	0.337	0.031	-0.249	-1.140
MAE		1.006	1.009	1.029	0.933	0.938	1.034	1.017	1.006	1.053	1.101	1.131	1.038
MSPE		0.983	0.988	1.066	0.891	0.942	1.069	1.051	1.051	1.117	1.186	1.243	1.077
6 Step Ahead													
Mean		-0.899	-0.758	-0.276	0.182	-3.144	2.375	1.304	-0.451	-0.089	-0.450	-0.997	-2.223
MAE		0.995	1.001	0.970	0.896	0.908	1.049	1.032	1.018	1.074	1.124	1.171	1.040
MSPE		0.972	0.980	1.100	0.880	0.931	1.120	1.086	1.087	1.195	1.280	1.370	1.139
12 Step Ahead													
Mean		-2.507	-2.246	-0.323	0.527	-5.959	2.643	1.402	-1.428	-1.158	-1.683	-2.494	-4.563
MAE		0.977	0.985	1.029	0.966	1.041	1.058	1.037	1.018	1.128	1.151	1.196	1.061
MSPE		0.967	0.978	1.170	0.951	1.228	1.150	1.135	1.094	1.313	1.440	1.401	1.171
			End Point Correction				**Post-Break Data**						
1 Step Ahead													
Mean	-0.368	-0.315	-0.397	-0.246	0.065		1.141	-0.120	-0.557			0.113	-1.860
MAE	0.966	0.961	0.960	1.061	0.975		1.030	0.996	0.990			1.105	1.075
MSPE	1.008	0.993	0.993	1.156	1.015		1.019	1.012	1.018			1.153	1.126
3 Step Ahead													
Mean	-0.273	-0.203	-0.306	-0.102	0.256		0.943	-0.372	-0.868			-0.030	-2.327
MAE	0.877	0.870	0.871	1.094	0.989		1.040	0.982	0.981			1.142	1.102
MSPE	0.856	0.830	0.830	1.214	0.951		1.057	1.007	1.027			1.319	1.258
6 Step Ahead													
Mean	-0.484	-0.324	-0.531	-0.088	0.603		0.597	-0.955	-1.612			-0.549	-3.846
MAE	0.731	0.725	0.721	1.040	0.905		1.049	0.987	0.983			1.195	1.146
MSPE	0.639	0.606	0.604	1.240	0.882		1.109	1.026	1.046			1.446	1.371
12 Step Ahead													
Mean	0.365	0.605	0.325	0.922	1.826		-0.755	-2.446	-3.414			-2.058	-7.768
MAE	0.595	0.572	0.570	1.033	0.842		1.050	0.994	0.991			1.194	1.186
MSPE	0.382	0.359	0.358	1.217	0.777		1.146	1.074	1.065			1.509	1.441

UK

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		0.713	0.713	-0.046	-0.043	2.197	2.796	2.163	1.991	0.415	0.183	1.223	0.713
MAE		1.005	1.005	1.024	0.995	0.997	1.000	1.008	1.018	1.002	1.000	0.998	1.006
MSPE		1.008	1.008	1.047	0.991	1.012	1.021	1.027	1.036	1.007	0.995	1.008	1.013
3 Step Ahead													
Mean		0.760	0.760	0.077	0.058	2.051	2.567	2.086	1.916	0.462	0.234	1.207	0.735
MAE		0.999	0.999	1.052	1.018	1.018	1.051	1.052	1.073	1.022	1.003	1.009	1.017
MSPE		1.006	1.006	1.117	1.047	1.049	1.096	1.096	1.141	1.044	0.992	1.028	1.039
6 Step Ahead													
Mean		0.765	0.765	0.056	0.041	2.070	2.583	2.099	1.904	0.443	0.226	1.209	0.737
MAE		0.997	0.997	1.090	1.051	1.069	1.123	1.124	1.160	1.059	0.999	1.026	1.033
MSPE		1.005	1.005	1.197	1.120	1.114	1.199	1.222	1.312	1.112	0.990	1.065	1.081
12 Step Ahead													
Mean		0.754	0.754	0.003	-0.009	2.104	2.596	2.097	1.863	0.395	0.209	1.198	0.736
MAE		0.974	0.974	1.125	1.089	1.106	1.189	1.217	1.278	1.062	0.951	1.069	1.043
MSPE		1.015	1.015	1.365	1.298	1.198	1.344	1.392	1.498	1.196	0.972	1.124	1.137
			End Point Correction					**Post-Break Data**					
1 Step Ahead													
Mean	-0.328	-0.394	-0.394	-0.410	-0.419		2.377	1.874	1.865			1.032	1.173
MAE	1.035	1.045	1.045	1.057	1.028		0.999	1.006	1.008			0.992	0.996
MSPE	1.088	1.099	1.099	1.147	1.084		1.012	1.021	1.009			0.975	0.977
3 Step Ahead													
Mean	-0.239	-0.291	-0.291	-0.279	-0.294		2.210	1.848	1.823			0.997	1.124
MAE	1.028	1.032	1.032	1.093	1.063		1.024	1.021	1.030			1.013	1.011
MSPE	1.078	1.088	1.088	1.238	1.156		1.058	1.053	1.067			1.004	1.004
6 Step Ahead													
Mean	-0.288	-0.341	-0.341	-0.318	-0.328		2.251	1.883	1.843			0.870	0.972
MAE	1.008	1.012	1.012	1.143	1.109		1.072	1.076	1.091			1.091	1.101
MSPE	1.034	1.044	1.044	1.328	1.244		1.114	1.123	1.159			1.222	1.234
12 Step Ahead													
Mean	-0.311	-0.366	-0.366	-0.334	-0.341		2.317	1.925	1.859			0.699	0.745
MAE	0.923	0.929	0.929	1.215	1.170		1.118	1.165	1.195			1.193	1.257
MSPE	0.912	0.939	0.939	1.503	1.426		1.196	1.241	1.302			1.620	1.724

GERMANY

Horizon	Linear	Diff	Pretest	2nd Diff	Smooth	AP	AP(c)	PSeq	PBIC	Seq	BIC	NPSeq	NPBIC
								Full Data Set					
1 Step Ahead													
Mean		0.407	0.447	0.021	0.425	1.017	0.214	-0.514	0.468	0.395	0.761	0.634	-0.119
MAE		0.942	0.950	1.130	0.975	0.965	1.019	1.139	0.998	1.104	1.000	1.045	1.100
MSPE		0.975	0.979	1.082	0.993	0.956	0.986	1.063	1.004	1.058	0.993	1.029	1.017
3 Step Ahead													
Mean		0.418	0.457	0.061	0.424	1.025	0.237	-0.465	0.486	0.404	0.774	0.639	-0.076
MAE		0.967	0.972	1.093	0.979	0.928	1.026	1.142	1.015	1.118	1.009	1.050	1.107
MSPE		0.937	0.943	1.074	0.950	0.910	0.956	1.090	1.001	1.121	0.981	1.064	1.030
6 Step Ahead													
Mean		0.430	0.469	0.074	0.433	1.058	0.257	-0.421	0.499	0.412	0.781	0.650	-0.037
MAE		0.959	0.965	1.098	0.966	0.939	1.018	1.130	1.026	1.087	1.020	1.055	1.100
MSPE		0.907	0.915	1.135	0.914	0.911	0.923	1.071	0.995	1.143	0.971	1.080	1.020
12 Step Ahead													
Mean		0.466	0.503	0.132	0.468	1.126	0.310	-0.296	0.537	0.447	0.798	0.683	0.063
MAE		0.968	0.976	1.125	0.970	0.981	1.019	1.109	1.023	1.056	1.024	1.038	1.073
MSPE		0.874	0.885	1.335	0.878	0.981	0.881	1.041	0.972	1.125	0.961	1.068	0.973
			End Point Correction					**Post-Break Data**					
1 Step Ahead													
Mean	-0.002	0.019	0.017	-0.051	0.019		0.095	-0.130	0.346			0.525	-0.164
MAE	1.014	1.013	1.013	1.333	1.035		1.012	1.125	0.994			1.021	1.076
MSPE	0.971	0.971	0.972	1.244	0.983		1.149	1.222	1.133			1.012	1.137
3 Step Ahead													
Mean	-0.008	0.013	0.012	-0.041	0.013		0.089	-0.129	0.368			0.531	-0.192
MAE	0.946	0.944	0.944	1.285	0.950		0.998	1.152	0.990			1.018	1.103
MSPE	0.806	0.805	0.806	1.306	0.813		0.997	1.179	1.030			1.025	1.086
6 Step Ahead													
Mean	-0.001	0.021	0.019	-0.023	0.021		0.085	-0.086	0.372			0.545	-0.172
MAE	0.853	0.849	0.850	1.261	0.851		0.979	1.118	1.002			1.016	1.090
MSPE	0.673	0.670	0.671	1.398	0.673		0.901	1.096	0.967			1.040	1.079
12 Step Ahead													
Mean	0.049	0.069	0.067	0.038	0.069		0.131	0.052	0.410			0.586	-0.049
MAE	0.651	0.647	0.648	1.142	0.648		0.996	1.095	1.022			1.031	1.090
MSPE	0.474	0.470	0.471	1.435	0.469		0.865	1.135	0.956			1.062	1.062

References

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with an Unknown Change Point. *Econometrica 61*, 821-856.

Andrews, D. W. K. and Ploberger, W. (1994). Optimal Tests when a Nuisance Parameter is Present Only under the Alternative. *Econometrica 62*, 1383-1414.

Bai, J. and Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica 66,* 47-78.

Bai, J. and Perron, P. (2001). Multiple Structural Change Models: A Simulation Analysis, forthcoming in *Econometric Essays*, D. Corbea, Durlauf, S. and Hansen, B. E. (eds.). Cambridge: University Press.

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

Blanchard, O. J. and Summers, L. H. (1987). Hysteresis and the European Unemployment Problem. NBER Working Paper No. W1950.

Chong, Y. and Hendry, D. (1986). Econometric Evaluation of Linear Macro-economic Models. *Review of Economic Studies 53*, 671 – 90.

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

Clements, M. and Hendry, D. (1999). *Forecasting Non-stationary economic time series*. The MIT Press: Cambridge, MA.

Diebold, F. and Kilian, L. (2000). Unit Root Tests are Useful for Selecting Forecasting Models. *Journal of Business and Economics Statistics 18*, 265-273.

Diebold, F. X. and Chen, C. (1996). Testing Structural Stability with Endogeneous Break Point: A Size Comparison of Analytic and Bootstrap Procedures. *Journal of Econometrics 70*, 221-241.

Elliott, G. (2004). Forecasting When There is a Single Break. Working paper, UCLA.

Enders, W. (2004). *Applied Econometric Time Series*. 2nd edition, John Wiley and Sons: Hoboken, N.J.

Enders, W. and Granger, C. (1998). Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates. *Journal of Business and Economic Statistics 16*, 304-11.

Koop, G, Pesaran, M. H. and Potter, S. (1996). Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics 74*, 119–147.

Pesaran, M. H. and Timmermann, A. (2005). Small Sample Properties of Forecasts from Autoregressive Models under Structural Breaks. *Journal of Econometrics 127*, 183-217.

Pesaran, M. H. and Timmermann, A. (2004). How Costly is to Ignore Breaks when Forecasting the Direction of a Time Series? *International Journal of Forecasting 20*, 411-424.

Phelps, E.S. (1994), "Low-wage Employment Subsidies versus the Welfare State" *American Economic Review: Papers and Proceedings, 84,* 54-58.

Prodan, R. (2006). Potential Pitfalls in Determining Multiple Structural Changes with an Application to Purchasing Power Parity. *Journal of Business and Economics Statistics*, forthcoming.

Rothman, P. (1998). Forecasting Asymmetric Unemployment Rates. *Review of Economics and Statistics 80*, 164 – 68.

Figure 1: A Persistent Series with Two Breaks

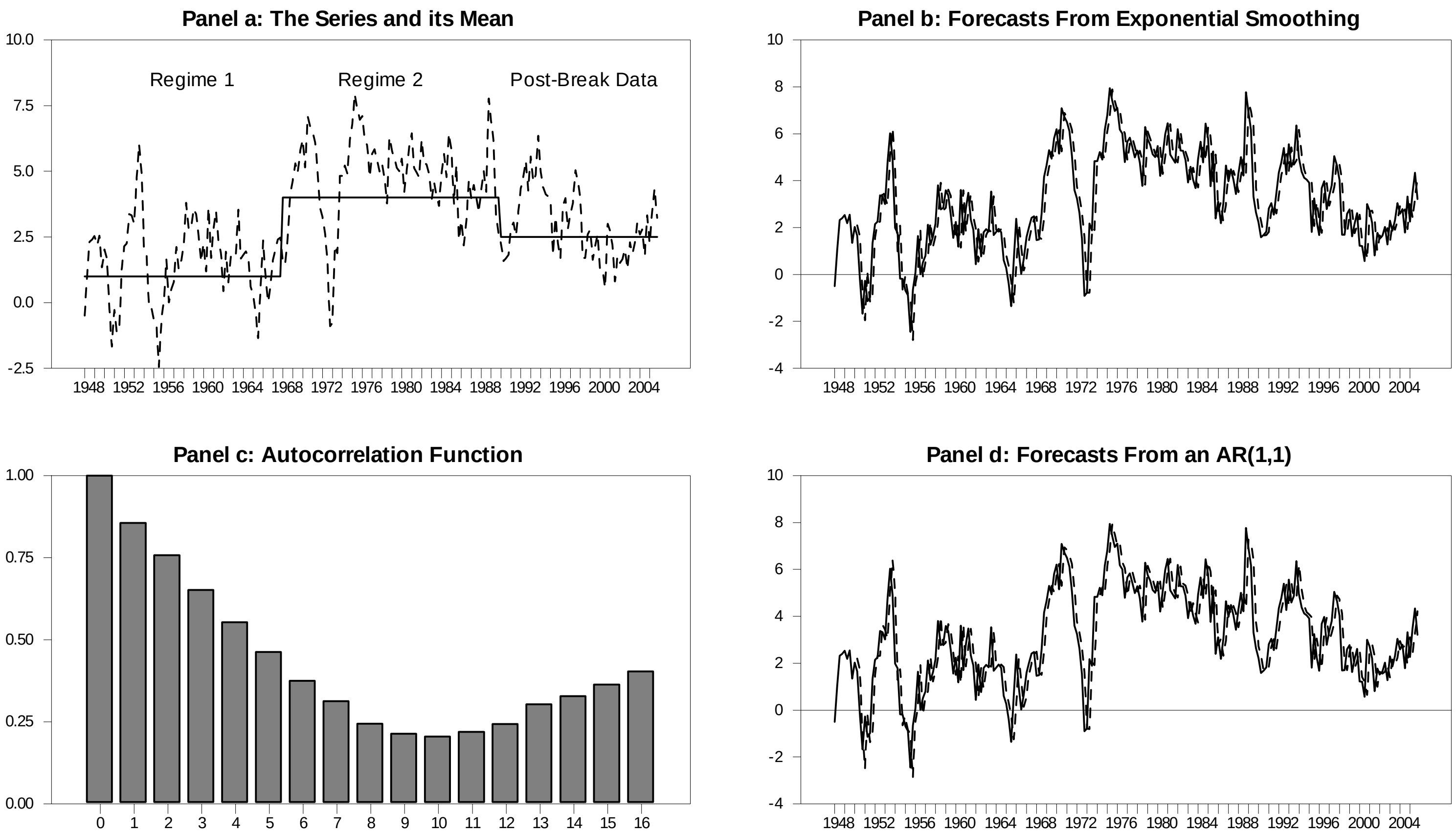

Figure 2a: Monthly Unemployment Rates

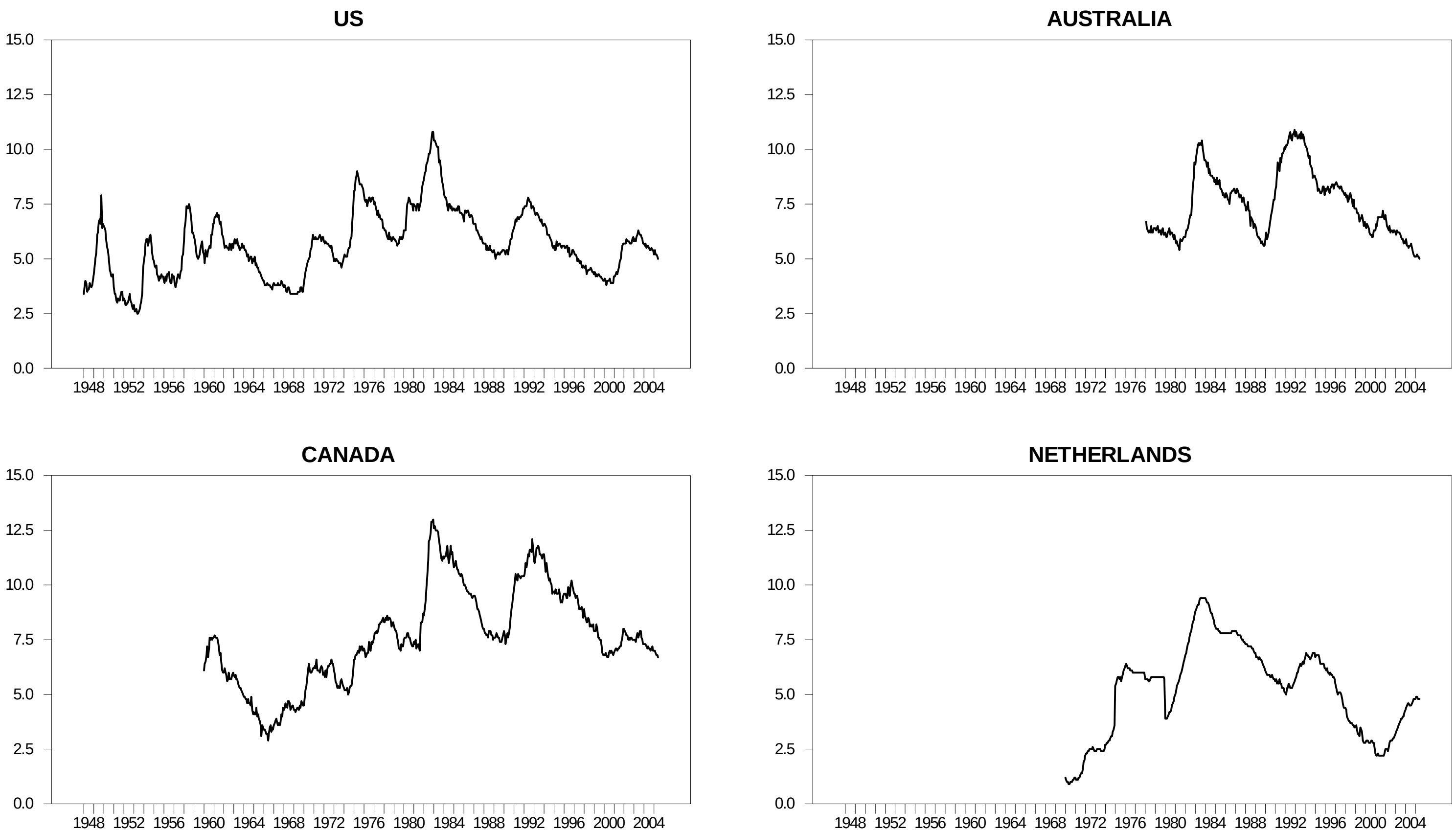

Figure 2b: Monthly Unemployment Rates

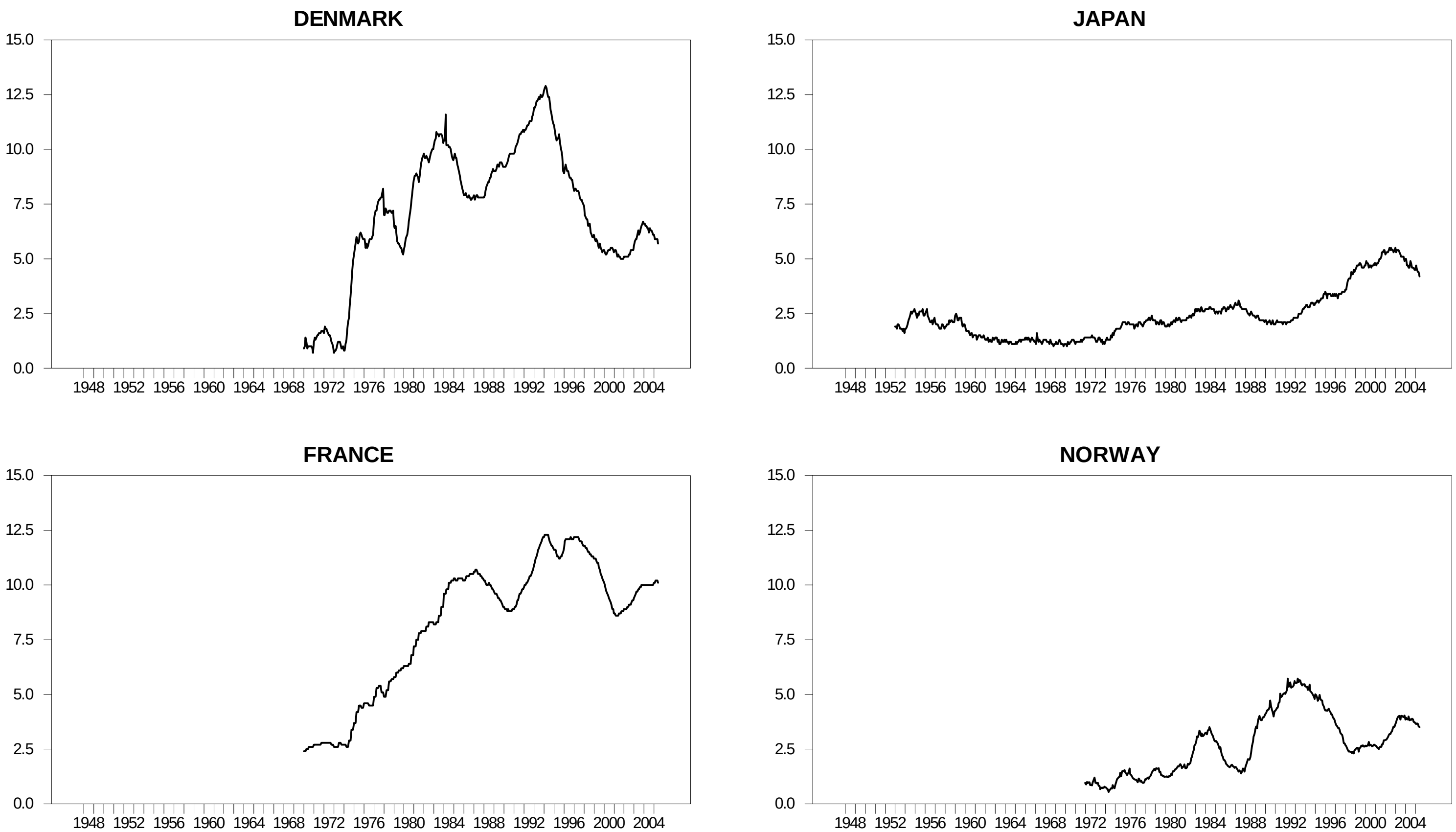

Figure 2c: Monthly Unemployment Rates

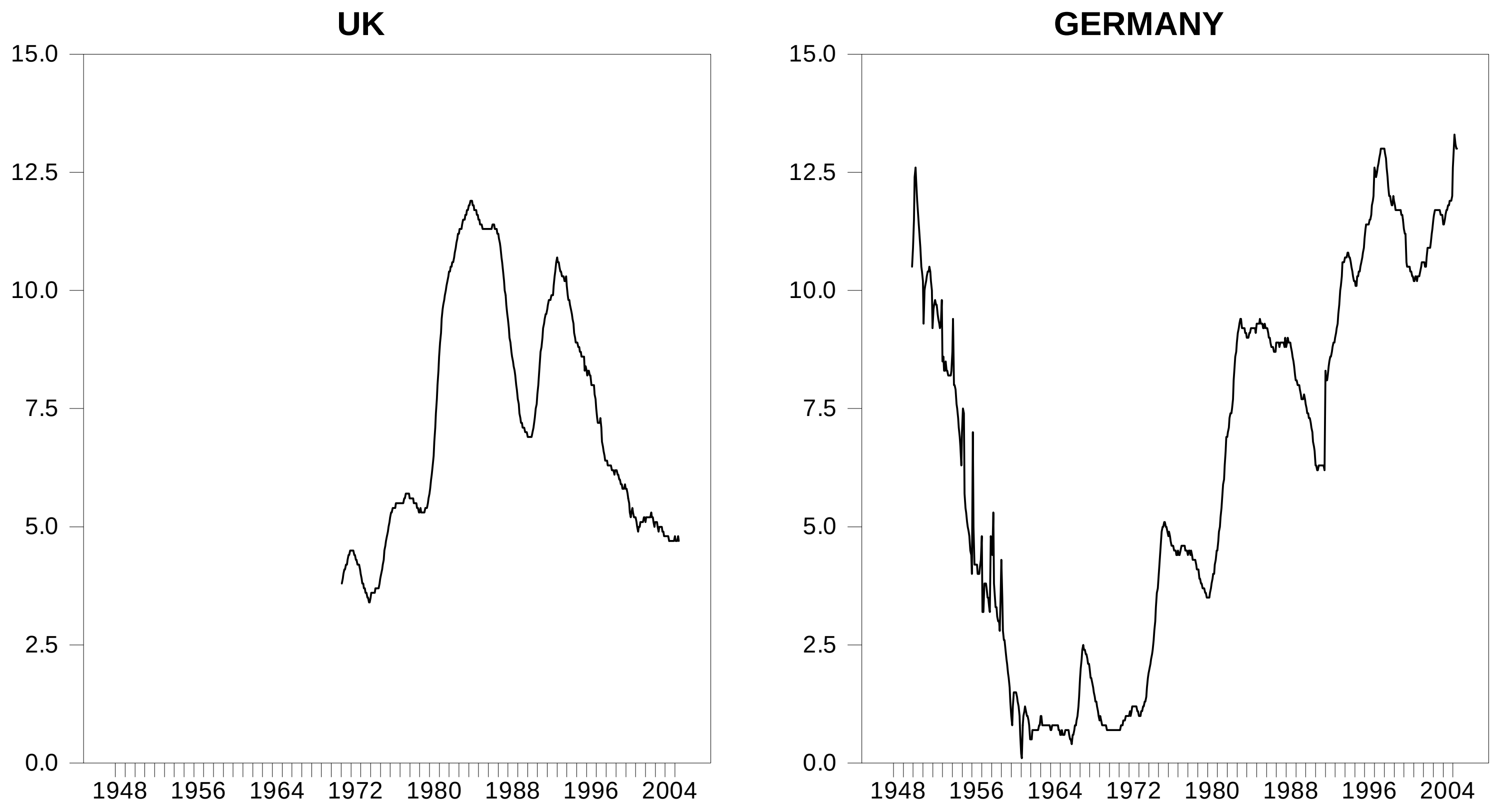

Table 1: Sub-sample Properties of the Data

		Sample Means				12-Lag Autocorrelations			
		All	1st third	2nd third	Last 3rd	All	1st third	2nd third	Last 3rd
	Start								
US	1948:1	5.63	4.82	6.44	5.63	0.699	0.286	0.752	0.745
CANADA	1960:1	7.63	5.28	8.79	8.78	0.880	0.680	0.698	0.808
AUSTRALIA	1978:2	7.54	7.46	8.39	6.78	0.735	0.594	0.669	0.826
NETHERLANDS	1970:1	5.22	4.12	7.15	4.40	0.900	0.819	0.760	0.885
DENMARK	1970:1	7.16	4.54	9.61	7.38	0.920	0.828	0.760	0.907
FRANCE	1970:1	8.23	4.36	9.67	10.66	0.973	0.919	0.662	0.840
JAPAN	1953:1	2.43	1.63	2.07	3.57	0.963	0.806	0.906	0.946
NORWAY	1972:1	2.76	1.29	3.53	3.48	0.913	0.509	0.858	0.797
UK	1971:2	7.29	5.61	9.91	6.40	0.929	0.835	0.790	0.923
GERMANY	1949:12	6.27	4.12	4.51	10.17	0.966	0.953	0.944	0.850

Table 2. The best forecasting model using an expanding window

a) 1 step

	MSPE	**MAE**	**Bias**
US	Linear	Linear	2nd difference
CANADA	Exp. smoothing	Exp. smoothing	2nd difference
AUSTRALIA	PBIC	Exp. smoothing	2nd difference
NETHERLANDS	Exp. smoothing	Exp. smoothing	Exp. smoothing
NORWAY	Exp. smoothing	Exp. smoothing	Exp. smoothing
DENMARK	1st difference	Exp. smoothing	2nd difference
FRANCE	Exp. smoothing	2nd difference	Exp. smoothing
GERMANY	1st difference	1st difference	2nd difference
JAPAN	2nd difference	2nd difference	PBIC
UK	Exp. smoothing	Exp. smoothing	Exp. smoothing

b) 3 steps

	MSPE	**MAE**	**Bias**
US	Linear	Linear	2nd difference
CANADA	Linear	1st difference	2nd difference
AUSTRALIA	Linear	Linear	2nd difference
NETHERLANDS	Exp. smoothing	Exp. smoothing	Exp. smoothing
NORWAY	Exp. smoothing	Exp. smoothing	AllBIC
DENMARK	Exp. smoothing	1st difference	2nd difference
FRANCE	AP	2nd difference	2nd difference
GERMANY	1st difference	1st difference	2nd difference
JAPAN	2nd difference	2nd difference	AP
UK	AllBIC	1st difference	Exp. smoothing

c) 6 steps

	MSPE	**MAE**	**Bias**
US	Linear	Linear	2nd difference
CANADA	Linear	Linear	2nd difference
AUSTRALIA	PSeq	1st difference	2nd difference
NETHERLANDS	Exp. Smoothing	Exp. smoothing	Exp. smoothing
NORWAY	Exp. Smoothing	Exp. smoothing	ALL sequential
DENMARK	1st difference	1st difference	2nd difference
FRANCE	AP	2nd difference	2nd difference
GERMANY	1st difference	1st difference	NoBIC
JAPAN	2nd difference	2nd difference	AP
UK	AllBIC	1st difference	Exp. smoothing

d) 12 steps

	MSPE	**MAE**	**Bias**
US	Linear	Linear	Exp. smoothing
CANADA	Linear	Linear	2nd difference
AUSTRALIA	1st difference	1st difference	2nd difference
NETHERLANDS	Exp. Smoothing	Exp. smoothing	Exp. smoothing
NORWAY	Exp. Smoothing	Exp. smoothing	2nd difference
DENMARK	1st difference	1st difference	2nd difference
FRANCE	PBIC	PBIC	2nd difference
GERMANY	1st difference	1st difference	NoBIC
JAPAN	AP	AP	PBIC
UK	AllBIC	AllBIC	2nd difference

Note: This table summarizes the models that have the smallest mean forecast error (**Bias**), mean absolute errors (**MAE**) and mean square prediction errors (**MSPE**) against the linear model at 1, 3, 6 and 12 steps ahead, for each country (except the nonlinear model). The competing models are: Linear, 1st difference, 2nd difference, Exponential smoothing, Andrew Ploberger, the partial-sequential (PSeq), the pure sequential (ALLSeq), the partial BIC (PBIC), the pure-BIC (AllBIC), the *non-parametric* sequential (NoSeq) and the *non-parametric* BIC (NoBIC).

Table 3. The TAR and MTAR models versus the linear model*

a) 1 step

	TAR			MTAR		
	Bias	**MAE**	**MSPE**	**Bias**	**MAE**	**MSPE**
US	**0.8902**	1.0344	1.0859	-3.1115	1.0952	1.2384
CANADA	6.2881	1.0776	1.1214	-7.0944	1.0189	1.0359
AUSTRALIA	2.3298	1.0997	1.2637	-2.9803	1.0613	1.2518
NETHERLANDS	1.1612	1.1425	1.3124	1.5430	1.3510	1.6215
NORWAY	3.0604	1.9250	37.5039	4.1685	1.0366	1.0142
DENMARK	2.2451	1.1275	1.3103	3.2463	1.1269	1.2187
FRANCE	**0.8436**	1.0496	1.0526	-3.1117	1.3026	1.7613
GERMANY	**0.8226**	1.0546	1.0691	0.9920	1.3768	1.2954
JAPAN	**0.9103**	1.0107	1.0099	0.5310	1.0032	1.0220
UK	2.1898	1.0200	1.0739	-2.5753	1.2955	1.7350

b) 3 steps

	TAR			MTAR		
	Bias	**MAE**	**MSPE**	**Bias**	**MAE**	**MSPE**
US	**0.8473**	1.0219	1.0485	-5.6155	1.2144	1.6545
CANADA	7.1233	1.1540	1.3108	-9.1618	1.0564	1.1300
AUSTRALIA	2.3211	1.1903	1.4315	**-0.8538**	1.1394	1.3125
NETHERLANDS	1.1494	1.1546	1.4541	**0.7578**	1.1180	1.1611
NORWAY	169.979	24.1073	38837.4	4.6509	1.0616	1.0469
DENMARK	2.1021	1.1775	1.3845	2.9905	1.1251	1.2394
FRANCE	1.0066	1.0929	1.1309	-3.7501	1.6047	3.7958
GERMANY	**0.8118**	1.0797	1.1298	1.1837	1.2266	1.2898
JAPAN	**0.9279**	1.0287	1.0557	**0.1234**	1.0427	1.3317
UK	2.3926	1.0788	1.2050	-2.4811	1.3891	2.2194

c) 6 steps

	TAR			MTAR		
	Bias	**MAE**	**MSPE**	**Bias**	**MAE**	**MSPE**
US	1.0317	1.0525	1.1003	-7.5245	1.4529	2.3555
CANADA	7.1986	1.2600	1.4682	-10.4850	1.0899	1.2499
AUSTRALIA	2.0782	1.3021	1.8567	**0.5920**	1.2526	1.6447
NETHERLANDS	1.0603	1.1115	1.3815	**0.5137**	1.0311	1.0266
NORWAY	100951	6802.30	6626128	6.8003	1.0555	1.1060
DENMARK	1.7836	1.2059	1.4561	3.1057	1.1904	1.4112
FRANCE	**0.9769**	1.0613	1.1329	-2.3350	1.7912	7.9648
GERMANY	**0.8351**	1.0891	1.1957	1.3015	1.1110	1.1819
JAPAN	1.2086	1.1243	1.4213	**-0.9629**	1.1624	5.1958
UK	2.4695	1.1594	1.3906	-1.8318	1.4083	2.3972

d) 12 steps

	TAR			MTAR		
	Bias	**MAE**	**MSPE**	**Bias**	**MAE**	**MSPE**
US	1.1359	1.0582	1.0934	-8.7343	1.9776	3.7140
CANADA	6.0899	1.3090	1.5597	-9.8333	1.1952	1.3968
AUSTRALIA	1.1064	1.3957	2.2929	1.0963	1.2620	1.6256
NETHERLANDS	1.0568	**0.9309**	**0.9815**	**0.4187**	**0.9486**	**0.8856**
NORWAY	3.51E+10	1.9E+09	4.91E+20	11.8828	1.2458	1.7184
DENMARK	1.5434	1.2318	1.5881	3.0619	1.2432	1.6006
FRANCE	**0.8727**	**0.9960**	1.1318	6.9287	3.7590	274.219
GERMANY	**0.8616**	1.1246	1.2734	1.3265	1.0411	1.1063
JAPAN	**0.9887**	1.0584	1.0923	-5.4887	1.5569	45.8431
UK	2.5483	1.3077	1.6162	-1.3006	1.3557	2.3521

Note: This table summarizes the mean forecast error (**Bias**), mean absolute errors (**MAE**) and mean square prediction errors (**MSPE**) for the **TAR** (threshold autoregressive) and **MTAR** (momentum threshold autoregressive) models relative to that of the linear model at 1, 3, 6 and 12 steps ahead for each country.

Table 4. The *t*-statistic for the Encompassing Test

Country	Horizon	PSeq	PBIC	AllSeq	AllBIC	NoSeq	NoBic
US	1-step	-3.86	**	-6.59	**	-3.93	-5.18
	3-step	-9.00	**	-15.25	**	-8.51	-12.45
	6-step	-13.74	**	-23.50	**	-12.47	-21.75
	12-step	-19.43	**	-13.63	**	-14.36	-32.51
CANADA	1-step	-1.47	**	-2.54	**	-2.53	-5.50
	3-step	-3.78	**	-4.87	**	-4.22	-10.31
	6-step	-5.36	**	-7.09	**	-5.72	-15.39
	12-step	-8.12	**	-9.48	**	-6.16	-19.06
AUSTRALIA	1-step	-1.59	0.47	-0.36	**	-3.07	-2.90
	3-step	-4.62	-1.29	-2.23	**	-6.98	-4.82
	6-step	-5.38	-4.79	-4.18	**	-11.86	-8.73
	12-step	-11.16	-8.77	-6.54	**	-13.13	-10.65
NETHERLANDS	1-step	-2.10	**	0.86	-8.89	-1.70	-1.53
	3-step	-3.20	**	0.78	-16.91	-2.94	-2.62
	6-step	-4.49	**	-1.35	-19.45	-5.42	-4.60
	12-step	-5.39	**	-4.31	-21.01	-7.41	-6.21
DENMARK	1-step	-4.45	-5.79	-1.84	-2.43	-4.21	-4.41
	3-step	-7.42	-8.50	-4.21	-4.65	-7.09	-7.62
	6-step	-9.61	-9.29	-6.31	-6.26	-9.73	-12.72
	12-step	-11.67	-9.83	-9.37	-8.68	-14.03	-19.19
FRANCE	1-step	-3.09	-0.49	-2.28	**	-1.76	-2.66
	3-step	-5.24	-0.73	-3.38	**	-2.41	-4.37
	6-step	-7.75	-0.88	-4.66	**	-2.96	-5.19
	12-step	-9.44	0.21	-6.08	**	-3.50	-6.40
JAPAN	1-step	2.41	-4.00	**	**	-0.59	-6.69
	3-step	5.26	-5.90	**	**	-0.36	-10.65
	6-step	6.75	-9.03	**	**	-0.66	-13.48
	12-step	8.80	-7.14	**	**	-0.32	-10.86
NORWAY	1-step	-2.23	-1.65	-3.59	-5.49	-3.86	-1.53
	3-step	-3.38	-2.83	-4.57	-7.58	-7.25	-2.82
	6-step	-4.39	-3.50	-5.47	-8.10	-9.38	-4.25
	12-step	-5.13	-3.93	-7.42	-8.81	-11.07	-6.35
UK	1-step	-2.77	-3.53	-1.43	-0.20	-1.29	-2.18
	3-step	-5.75	-7.30	-3.23	-0.68	-2.49	-3.60
	6-step	-8.05	-10.26	-5.13	-1.52	-3.64	-4.67
	12-step	-10.27	-13.54	-7.35	-2.83	-5.52	-6.14
GERMANY	1-step	-6.17	-1.99	-5.61	0.12	-4.27	-1.74
	3-step	-5.88	-1.84	-7.37	0.43	-5.70	-2.85
	6-step	-5.47	-1.65	-7.81	0.59	-6.36	-3.16
	12-step	-5.36	-0.97	-7.14	0.83	-5.74	-2.68

Note: For each variant of the BP model, entries are the *t*-statistic for the null hypothesis $\gamma = 0$.
** denotes that the BP test yielded no breaks so that the forecasts from the linear and BP model were identical.