

An Empirical Study of Seasonal Unit Roots in Forecasting

Michael P. Clements
Department of Economics,
University of Warwick

and

David F. Hendry*
Nuffield College,
Oxford.

March 31, 2008

Abstract

We assess the usefulness of pre-testing for seasonal roots, based on the HEGY approach, for out-of-sample forecasting. It is shown that if there are shifts in the deterministic seasonal components then the imposition of unit roots can partially robustify sequences of rolling forecasts, yielding improved forecast accuracy. The analysis is illustrated with two empirical examples where more accurate forecasts are obtained by imposing more roots than is warranted by HEGY.

The issue of assessing forecast accuracy when predictions of any one of a number of linear transformations may be of interest is also addressed.

1 Introduction

An important strand of the recent literature on modelling seasonality in economic time series has been concerned with testing for seasonal unit roots. Hyllberg, Engle, Granger and Yoo (1990) (henceforth, HEGY) helped popularize the modelling of economic time series as variables which exhibit seasonal unit roots, but as those authors acknowledge, time-series analysts in the tradition of Box and Jenkins (1970) implicitly assumed the existence of seasonal unit roots by using the seasonal differencing operator. However, there is little evidence in the literature on the impact of imposing seasonal unit roots on forecast accuracy. An exception is Paap, Franses and Hoek (1997), who compare the forecast performance of autoregressive seasonal unit root models and seasonal mean shift models, both empirically and by Monte Carlo simulation. A main focus of this paper is to assess the usefulness of pre-testing for seasonal roots, based on the HEGY approach, when the aim of the exercise is out-of-sample forecasting.

To this end, in section 2 we briefly discuss testing for seasonal roots within the HEGY framework, and the tension that is often apparent in practice between the number of unit roots implied by this approach compared to that of Box and Jenkins (1970).

Then, we return to some issues that have been of concern in our recent work. These relate to what are appropriate measures of forecast accuracy, and would appear to carry extra force in the context of forecasting seasonal series. In Clements and Hendry (1993b), we questioned the usefulness of some commonly-used measures of forecast accuracy for non-seasonal series, due to their lack of invariance under linear transformations for which the associated model class is invariant, and developed an alternative. In section 3 we elaborate on the view that these concerns would appear to be at least as relevant for seasonal series, and we briefly discuss an invariant measure of forecast accuracy.

*Financial support from the UK Economic and Social Research Council under grants R000233447 and L11625015 is gratefully acknowledged by both authors. The paper has benefited from helpful comments from Philip Hans Franses and three anonymous referees, to whom we wish to express our gratitude. The calculations reported in this paper were carried out using GAUSS-386i VM Version 3.2.6 and the graphs were produced using GiveWin 9.00 beta01, a pre-release Windows version of PcGive. A more detailed set of results than presented here is available on request from the authors.

In section 4 we consider the impact of structural breaks in seasonals, and show that the imposition of unit roots can partially robustify sequences of rolling forecasts.

In section 5 we illustrate the analyses with applications to forecasting income in Norway and consumption in the UK, and interpret the findings in light of the model of seasonal shifts. Section 6 concludes.

2 Non-periodic models of seasonality

We consider forecasting based on scalar autoregressive (AR) models of seasonal processes, and in particular the usefulness of imposing roots based upon pre-testing procedures. However, some doubt has been cast upon the reliability of tests for seasonal roots based on AR models given that empirically there often appears to be a tension between inference concerning the number of unit roots suggested by such tests, and the number implicit in the modelling approach of Box and Jenkins.

For quarterly data, the member of the general class of seasonal autoregressive-moving average (SARMA) models developed by Box and Jenkins (1970) that has proved successful empirically takes the form:

$$(1 - L)(1 - L^4)x_t = \mu + (1 - \theta_1 L)(1 - \theta_4 L^4)\epsilon_t \quad (1)$$

where $\epsilon_t \sim \text{IN}(0, \sigma\epsilon^2)$, $|\theta_1| < 1$, $|\theta_4| < 1$ and $L^k x_t = x_{t-k}$. Franses (1996) p.42–46 references empirical studies that fit SARMA models and provides examples. This model is sometimes known as the ‘airline’ model. The filter $\Delta_4 = (1 - L^4)$ relates to the seasonal part of the model, and captures the tendency for the value of the series in a particular season to be highly correlated with the value in the same season a year earlier. The filter $\Delta = 1 - L$ relates to the nonseasonal part of the model and captures the stochastic trend that appears to characterise many economic time series (for example, Nelson and Plosser, 1982, but other views exist: Perron, 1989, 1990). Thus (1) implies annual and one-period (zero-frequency) differencing filters in the autoregressive polynomial. Since the AR polynomial in (1) can be written as:

$$(1 - L)(1 - L^4) = (1 - L)[(1 - L)(1 + L)(1 - iL)(1 + iL)]$$

the airline model entails two zero-frequency roots, as well as roots at the bi-annual and seasonal frequencies (taking the two complex roots as $(1 - iL)(1 + iL) = (1 + L^2)$).

On the other hand, the HEGY procedure seldom finds more than a single zero-frequency root and only occasionally roots at all the seasonal frequencies (see, for example, Osborn, 1990, Hylleberg, Jørgensen and Sørensen, 1993, and Franses, 1996, Table 5.2, p.68).

A number of possible resolutions to this conundrum have been proposed. If (1) is correct, then the neglect of the MA components in the HEGY testing procedure may lead to test sizes exceeding their nominal values (see, for example, Franses and Koehler, 1994), analogous to findings in the nonseasonal unit-root testing literature (see, for example, Hall, 1989). Additional lags of the dependent variable can be included to capture the autocorrelation in the errors, but the size distortion is bound to be large for tests of nonseasonal and seasonal unit roots if there is near-cancellation with the MA roots.

Another possibility is that ‘over-differencing’ in (1) converts level shifts in seasonal means into ‘blips’ which may be mistaken for outliers. For example, suppose the data generating process (DGP) is:

$$y_t = \sum_{i=1}^4 (\alpha_i + \alpha_i^* \mathbf{I}_T^\tau) Q_{i,t} + \epsilon_t \quad (2)$$

that is, consists of a deterministic seasonal component, with a step change in the seasonal means when $\alpha_i^* \neq 0$, for some i , where $\mathbf{I}_T^\tau = 1$ for $t \in (\tau, T)$ and is zero otherwise. In practice, y_t could be $\Delta_1 x_t$,

for example. Thus, the seasonal means change at $t = \tau$. The $Q_{i,t}$ are seasonal dummy variables defined by $Q_{i,t} = 1$ if $t = i + 4q$, $q = 0, 1, 2, \dots$, and zero otherwise. Taking fourth differences would convert the step changes to blips:

$$\Delta_4 y_t = \left(\sum_{i=1}^4 \alpha_i \Delta_4 Q_{i,t} + \alpha_i^* \Delta_4 \mathbf{I}_T^\tau Q_{i,t} \right) + \Delta_4 \epsilon_t = \sum_{i=1}^4 \alpha_i^* \mathbf{I}_{\tau+3}^\tau Q_{i,t} + \Delta_4 \epsilon_t$$

This model has seasonal and zero-frequency unit roots, a (non-invertible) negative MA and four outliers caused by the change in the seasonal means. Omitting dummies for the seasonal-shift outliers, and not modelling the MA term might lead to the number of autoregressive unit roots being over-estimated. It is well known in the nonseasonal unit root testing literature that structural breaks may lead to the number of unit roots being over-estimated (for example, Perron, 1989, 1990, Hendry and Neale, 1991). Seasonal differencing of this form to convert level changes to blips in seasonal processes is the subject of section 4, although we do not in this paper apply the recent extensions to the HEGY procedure that have been developed to allow for shifts in seasonal means (see, for example, Franses and Vogelsang, 1997, Smith and Otero, 1997).

We close this section by briefly describing the HEGY procedure as implemented in this paper to test for seasonal unit roots. The procedure is based upon the regression equation:

$$\alpha(L) \Delta_4 x_t = \pi_1 z_{1,t-1} + \pi_2 z_{2,t-1} + \pi_3 z_{3,t-2} + \pi_4 z_{3,t-1} + \mu_t + \epsilon_t \quad (3)$$

where:

$$\mu_t = \mu_1 + \mu_2 t + \sum_{i=1}^3 \mu_{2+i} Q_{i,t},$$

$\alpha(L)$ is a stationary p^{th} order lag polynomial, and:

$$\begin{aligned} z_{1t} &= (1 + L + L^2 + L^3)x_t = S(L)x_t \\ z_{2t} &= -(1 - L)(1 + L^2)x_t \\ z_{3t} &= -(1 - L^2)x_t = -(1 - L)(1 + L)x_t. \end{aligned}$$

The tests for unit roots at different frequencies are based on the significance of the estimated $\hat{\pi}_i$ coefficients. For example, if $\hat{\pi}_1$ is significantly different from, typically being less than, zero, then the null of a unit root at the zero frequency is rejected. A rejection of the null that π_2 equals zero indicates the absence of a unit root at the bi-annual frequency. Roots at the annual frequency can be tested by an F -test of the null that $\pi_3 = \pi_4 = 0$. Typically the order p of the polynomial $\alpha(L)$ is chosen to make the residuals from (3) approximately white noise. Hylleberg *et al.* (1990) provide the details.

The HEGY procedure can also be adapted to generate models for forecasting seasonal processes, that emphasize the role played by imposing seasonal roots rather than implicitly estimating them. Notice that (3) with $\pi_i \neq 0$, $i = 1, \dots, 4$, is equivalent to an $\text{AR}(p + 4)$ model. We use the HEGY procedure to test for seasonal roots, as outlined above, where p is the smallest lag order consistent with non-rejection on an LM-test of serial correlation up to fourth order. Then, the regression is re-run omitting z_i which are not significantly different from zero (judged using the appropriate critical values: see Hylleberg *et al.*, 1990). All roots are imposed when the null hypotheses that $\pi_1 = 0$, $\pi_2 = 0$ and $\pi_3 = \pi_4 = 0$ are not rejected, which implies that the process is $\text{I}(1)$ (zero frequency root) after seasonally adjusting with the $S(L)$ filter. The deterministic trend is retained if its t -value exceeds 2 in the first regression, although the t -distribution may not be appropriate. This (restricted) estimated model then becomes the forecasting model.

3 Forecast evaluation criteria

Clements and Hendry (1995a) consider the importance of imposing zero-frequency unit roots in a bivariate nonseasonal cointegrated system, and show that the rankings (based on mean squared forecast error) of alternative methods partly depend on the linear transformation of the series for which predictions are evaluated. In terms of predicting either the first differences of the series or the stationary ($I(0)$) linear combination, the improvement in accuracy is small from imposing roots. For seasonal series, we may be interested in predicting any of the separate factors of the annual differencing operator:

$$(1 - L^4) = (1 - L)(1 + L)(1 + L^2) \quad (4)$$

as well as the annual percentage change (assuming the series are in logarithms). For example, the UK Treasury sometimes presents half-yearly forecasts (i.e. of $(1 + L)x_t$) based on models of quarterly data series. Sometimes the second and third factors on the right-hand side of (4) are taken together as:

$$S(L) \equiv (1 + L)(1 + L^2) = 1 + L + L^2 + L^3$$

to denote the seasonal filter, but we will find it useful to refer to the last factor in (4) as the seasonal filter given $(1 + L)x_t$ may be of interest separately.

Clements and Hendry (1993b) show that a model which yields better forecasts on one linear transformation than a rival may lose out on another when the metric for assessing forecast accuracy is (root) mean squared forecast error, (R)MSFE. Thus, model 1 may have a smaller MSFE for predicting $(1 - L)x_t$ than model 2 but a larger MSFE for predicting $(1 + L)x_t$ say, for multi-step forecasts. Hence, as well as calculating RMSFEs for linear transformations of the data, we also report generalized forecast-error second moment (GFESM) calculations, which are invariant under all non-singular scale-preserving linear transformations of the series.

The GFESM for h -step forecasts is the determinant of the outer-product of the vector formed by stacking the forecast errors from all previous step ahead forecasts ($1, 2$ up to h) as in (5):

$$|\Phi_h| = |E[\mathbf{E}_h \mathbf{E}_h']| \quad (5)$$

where $\mathbf{E}_h' = (e_{T+1}', e_{T+2}', \dots, e_{T+h-1}', e_{T+h}')$, where the e_{T+j} denote forecast errors, and are scalars for the univariate processes considered in this paper.

Before establishing the invariance of $|\Phi_h|$, we demonstrate the invariance of forecasts to linear transforms. Denote a univariate forecasting model by the succinct notation:

$$\phi' \mathbf{s}_t = u_t \quad (6)$$

where $u_t \sim \text{ID}(0, \sigma^2)$, $\mathbf{s}_t' = (x_t' : \mathbf{z}_t')$, x_t is the variable to be forecast and \mathbf{z}_t are k available predetermined variables (here just lags of x_{t-1}), and $\phi = (1 : -\beta')'$ say. The parameters are (β, σ^2) . Then the likelihood of (6) is invariant under non-singular transformations of the form:

$$\phi' \mathbf{P}^{-1} \mathbf{P} \mathbf{s}_t = u_t$$

which we can write as:

$$\phi^{*'} \mathbf{s}_t^* = u_t. \quad (7)$$

In (7), $\mathbf{s}_t^* = \mathbf{P} \mathbf{s}_t$, \mathbf{P} is a $(k + 1) \times (k + 1)$ non-singular matrix which is upper block-triangular, for example:

$$\mathbf{P} = \begin{pmatrix} 1 & \mathbf{P}_{12}' \\ \mathbf{0} & \mathbf{P}_{22} \end{pmatrix} \quad \text{so that} \quad \mathbf{P}^{-1} = \begin{pmatrix} 1 & -\mathbf{P}_{12}' \mathbf{P}_{22}^{-1} \\ \mathbf{0} & \mathbf{P}_{22}^{-1} \end{pmatrix}.$$

Since we need to calculate \mathbf{P}^{-1} , a restriction on \mathbf{P} is that $|\mathbf{P}_{22}| \neq 0$. Then:

$$\phi^{*'} \equiv \phi' \mathbf{P}^{-1} = (1 : -(\mathbf{p}_{12} + \beta)' \mathbf{P}_{22}^{-1}) = (1 : -\beta^{*'}) .$$

No restrictions are imposed by these transforms, so equations (6) and (7) are isomorphic. Transformations in the class shown in (7) include all the right-hand side factors of (4). For example, the bi-annual filter for an AR(2) model has $\mathbf{p}'_{12} = (1 \ 0)$, $\mathbf{P}_{22} = \mathbf{I}_2$, and $\mathbf{s}_t = (x_t \ x_{t-1} \ x_{t-2})$, $\beta' = (\beta_1 \ \beta_2)$, say. The formalization merely makes explicit the fact that a linear model is defined by its invariance under linear transformations (or affine transforms more generally). Forecasts and forecast confidence intervals made in the original system and transformed after the event to x_t^* , or made directly from the transformed system, are identical. This remains true when parameters are estimated by any method that is invariant (for example, maximum likelihood), and holds for multi-period forecasts. Hence forecasts are invariant to linear transforms: the critique in Clements and Hendry (1993b) is of the lack of invariance of some measures of their accuracy, such as certain MSFE-based measures.

For transformations involving \mathbf{P} , measures based on the ‘simple’ MSFE matrix $\mathbf{E}[e_{T+j} e'_{T+j}]$ are not invariant. It is necessary to account for covariances between different step-ahead errors, as in the GFESM. Suppose $\mathbf{P} = \Delta \mathbf{I}_n$, that is, a diagonal matrix of differencing operators. From the invariance of forecasts to linear transforms, we have that $\mathbf{e}_{\Delta, T+j|T} = \mathbf{e}_{T+j|T} - \mathbf{e}_{T+j-1|T}$, where $\mathbf{e}_{T+j|T}$ is the error in forecasting \mathbf{x}_{T+j} conditional on T , and $\mathbf{e}_{\Delta, T+j|T}$ is the error in forecasting $\Delta \mathbf{x}_{T+j}$. Hence writing $\mathbf{E}'_{\Delta, h} = (\mathbf{e}'_{\Delta, T+1|T}, \mathbf{e}'_{\Delta, T+2|T}, \dots, \mathbf{e}'_{\Delta, T+h-1|T}, \mathbf{e}'_{\Delta, T+h|T})$, we obtain:
 $\mathbf{E}_{\Delta, h'} = (\mathbf{e}'_{T+1} \mathbf{P}', \mathbf{e}'_{T+2} \mathbf{P}', \dots, \mathbf{e}'_{T+h-1} \mathbf{P}', \mathbf{e}'_{T+h} \mathbf{P}')$, or $\mathbf{E}_{h, \Delta} = (\mathbf{I} \otimes \mathbf{P}) \mathbf{E}_h$, thus:

$$|\Phi_{\Delta, h}| = |\mathbf{E} [\mathbf{E}_{h, \Delta} \mathbf{E}_{h, \Delta}']| = |\mathbf{E} [(\mathbf{I} \otimes \mathbf{P}) \mathbf{E}_h \mathbf{E}_h' (\mathbf{I} \otimes \mathbf{P}')]| = |\mathbf{E} [\mathbf{E}_h \mathbf{E}_h']| \times |\mathbf{I} \otimes \mathbf{P}|^2 = |\mathbf{E} [\mathbf{E}_h \mathbf{E}_h']|$$

since $|\mathbf{I} \otimes \mathbf{P}| = |\mathbf{I} \otimes \mathbf{P}'| = 1^h \times |\mathbf{P}|^h = 1$ and $\mathbf{P} \mathbf{e}_{T+1} = \mathbf{e}_{T+1}$.

We have pin-pointed lack of invariance as one advantage of GFESM over MSFE-based measures for evaluating the accuracy of predictions of seasonal series, given that in principle there are a variety of different transforms that may be of interest. There are of course a number of disadvantages that both measures share and some problems peculiar to the GFESM (see, for example, Ericsson, 1992, Clements and Hendry, 1998b, the discussants’ comments on our original paper, Clements and Hendry, 1993b, and the reply, Clements and Hendry, 1993a, and also Armstrong and Fildes, 1995 and Clements and Hendry, 1995b) and space precludes us reviewing those arguments here. However, a common remark is that invariance has little to recommend it if only a particular transformation is of interest, with which we concur. More fundamentally, it is suggested that econometric analyses should begin by writing down a loss function that maps forecast errors to their economic costs, with the optimal predictor being derived as the loss minimizer. In practice, well-defined mappings between forecast errors and costs are not typical (Granger, 1993, p.651: ‘real-world cost functions are rarely available’) – as are users who care about predicting only one linear transform.

A number of recent papers have considered the importance of matching the form of the model estimation criterion with that for forecast evaluation, following Granger (1969), and see also Zellner (1986). The evidence in Weiss and Andersen (1984) for the practical relevance of matching is not strong, and Fildes and Makridakis (1988) question its practicality (and even desirability, with reference to Nelson and Granger, 1979). More recently, Weiss (1996) takes up this suggestion and explores the properties of procedures which base estimation on the forecast criterion function, defined over l -step errors, for a variety of forecast criteria, such as asymmetric quadratic loss (also see Christoffersen and Diebold, 1997).

The models we use for forecasting are estimated by OLS. The OLS estimates are used to obtain forecasts for all future periods, so that no attempt is made to robustify the longer-horizon forecasts by estimating the model to minimize, say, the j -step ahead in-sample sum of squares, if the aim is to

generate j -step ahead ex ante forecasts. In principle there may be gains to multi-step estimation whenever the model is incorrectly specified. The literature began with Cox (1961), who considered making the smoothing parameter of the exponential smoothing formula dependent on the forecast horizon, an idea extended by Tiao and Xu (1993), with contributions by, *inter alia*, Weiss (1991), Stoica and Nehorai (1989), Tsay (1993), and Clements and Hendry (1996b). The latter paper indicates gains to multi-step estimation when the model omits an MA component with negative MA parameters close to -1 ; as would be the case for AR models of seasonal processes if the airline model applied. This may be an interesting avenue for future research, although the outcome remains dependent on the exact criterion selected for minimization, so very different decisions could result for levels, first or seasonal differences, and so on. Consequently, we did not pursue this route since all the transforms reported in section 5 are of interest to some users of macroeconomic forecasts.

4 The impact of structural breaks in seasonals

Consider a quarterly logarithmic-transformed time series described by fixed, trending and shifting deterministic seasonals generalizing (2):

$$y_t = \sum_{i=1}^4 (\alpha_i + \alpha_i^* \mathbf{I}_T \tau + \beta_i t + \beta_i^* \mathbf{I}_T \tau (t - \tau + 1)) Q_{i,t} + \epsilon_t \quad (8)$$

for $t = 1, \dots, T$, where the indicator variable $\mathbf{I}_{\tau+j} \tau = 1$ for $t \in (\tau, \tau + j)$ and is zero otherwise as above. We neglect additional sources of dynamics to focus on a change in the seasonal pattern at time τ . The issue we wish to address here concerns conditions under which $\Delta_4 y_t$ has a constant, non-zero seasonal pattern or a changing seasonal pattern, to help account for the later empirical results which reveal that it pays to impose unit and/or seasonal roots even when it is statistically invalid to do so. As above, we assume that $\epsilon_t \sim \text{IN}(0, \sigma \epsilon^2)$.

From (8), in terms of annual changes, $\Delta_4 = (1 - L^4)$:

$$\begin{aligned} \Delta_4 y_t &= \sum_{i=1}^4 (\alpha_i \Delta_4 Q_{i,t} + \alpha_i^* \Delta_4 \mathbf{I}_T \tau Q_{i,t} + \beta_i \Delta_4 t Q_{i,t} + \beta_i^* \Delta_4 \mathbf{I}_T \tau (t - \tau + 1) Q_{i,t}) + \Delta_4 \epsilon_t \\ &= \sum_{i=1}^4 ((\alpha_i^* + \beta_i^* (t - \tau + 1)) \mathbf{I}_{\tau+3} \tau Q_{i,t} + 4\beta_i Q_{i,t} + 4\beta_i^* \mathbf{I}_T^{\tau+4} Q_{i,t}) + \Delta_4 \epsilon_t, \end{aligned} \quad (9)$$

since $Q_{i,t} \equiv L^4 Q_{i,t}$ so $\Delta_4 Q_{i,t} = 0 \ \forall i, t$; $\Delta_4 t Q_{i,t} = (t Q_{i,t} - L^4 t Q_{i,t}) = 4Q_{i,t}$; $\Delta_4 \mathbf{I}_T \tau Q_{i,t} = \mathbf{I}_{\tau+3} \tau Q_{i,t}$; and $\Delta_4 \mathbf{I}_T \tau (t - \tau + 1) Q_{i,t} = (t - \tau + 1) \mathbf{I}_{\tau+3} \tau Q_{i,t}$ for $t - \tau = 0, 3$ and $4\mathbf{I}_T^{\tau+4} Q_{i,t}$ for $t - \tau > 3$. When the existence and timing of shifts in seasonality are unknown, all the terms involving products with the indicator variable will add to the residual, so over the sample as a whole, there will not be a redundant common factor of $(1 - L^4)$. The residuals on (9) are likely to be negatively autocorrelated in the absence of any dynamic modelling, offset by any original positive autocorrelation in the $\{\epsilon_t\}$.

In the neighbourhood of time τ , the evolving seasonal patterns of y_t and $\Delta_4 y_t$ are as follows.

t	$E[y_t]$	$E[\Delta_4 y_t]$	
\vdots	\vdots	\vdots	
$\tau - 1$	$\sum_{i=1}^4 (\alpha_i + \beta_i (\tau - 1)) Q_{i,\tau-1}$	$4 \sum_{i=1}^4 \beta_i Q_{i,\tau-1}$	
τ	$\sum_{i=1}^4 (\alpha_i + \alpha_i^* + \beta_i \tau + \beta_i^*) Q_{i,\tau}$	$\sum_{i=1}^4 (4\beta_i + \alpha_i^* + \beta_i^*) Q_{i,\tau}$	
$\tau + 1$	$\sum_{i=1}^4 (\alpha_i + \alpha_i^* + \beta_i (\tau + 1) + 2\beta_i^*) Q_{i,\tau+1}$	$\sum_{i=1}^4 (4\beta_i + \alpha_i^* + 2\beta_i^*) Q_{i,\tau+1}$	(10)
$\tau + 2$	$\sum_{i=1}^4 (\alpha_i + \alpha_i^* + \beta_i (\tau + 2) + 3\beta_i^*) Q_{i,\tau+2}$	$\sum_{i=1}^4 (4\beta_i + \alpha_i^* + 3\beta_i^*) Q_{i,\tau+2}$	
$\tau + 3$	$\sum_{i=1}^4 (\alpha_i + \alpha_i^* + \beta_i (\tau + 3) + 4\beta_i^*) Q_{i,\tau+3}$	$\sum_{i=1}^4 (4\beta_i + \alpha_i^* + 4\beta_i^*) Q_{i,\tau+3}$	
$\tau + 4$	$\sum_{i=1}^4 (\alpha_i + \alpha_i^* + \beta_i (\tau + 4) + 5\beta_i^*) Q_{i,\tau+4}$	$4 \sum_{i=1}^4 (\beta_i + \beta_i^*) Q_{i,\tau+4}$	
$\tau + 5$	$\sum_{i=1}^4 (\alpha_i + \alpha_i^* + \beta_i (\tau + 5) + 6\beta_i^*) Q_{i,\tau+5}$	$4 \sum_{i=1}^4 (\beta_i + \beta_i^*) Q_{i,\tau+5}$	
\vdots	\vdots	\vdots	

First, consider the pattern in y_t : the initially steady, if trending, seasonality is disturbed at τ , with a jump in the level when $\alpha_i^* \neq 0$ and in the trend when $\beta_i^* \neq 0$. The new level in each quarter becomes $\alpha_i + \alpha_i^* + \beta_i \tau$ and the new trend is $(\beta_i + \beta_i^*)$, so forecasts based on assuming $\alpha_i^* = \beta_i^* = 0$ will quickly become very poor.

Next, consider $\Delta_4 y_t$: when there are no trending seasonals ($\beta_i = \beta_i^* = 0$), the break just produces a jump in the quarterly change in the annual growth rate equal to α_i^* for the four quarters from τ to $\tau + 3$. When there are no structural breaks, $\Delta_4 y_t$ has a constant seasonal pattern equal to $4\beta_i$, and so systematically changes more in some quarters than others. Also, there are complicated movements in the seasonal pattern of $\Delta_4 y_t$ whenever the coefficients α_i^* and β_i^* differ from zero (i.e., $I_T \tau = 1$). Finally, there is a permanent shift in the seasonal pattern when $\beta_i^* \neq 0$. Analogous to differencing a non-seasonal process, the Δ_4 filter reduces the impact of the effect of α_i^* to a transient ‘blip’ for four quarters, and that of β_i^* from a trend to a level shift.

Given the above patterns induced by changes in deterministic seasonals, we can show that imposing unit roots in the forecast model may improve the accuracy of rolling forecasts when the DGP is (8) and the shifts are not modelled. The advantages to using, say, fourth-differencing, arise for the same reason that zero-frequency differencing works in the nonseasonal case when there are structural breaks or regime shifts: see Clements and Hendry (1996a, 1998a). Essentially, differencing sets the forecasts ‘back on track’ after the break has occurred. The ‘full value’ of the break is incorporated in future forecasts since differencing effectively projects only the present and gives no weight to the pre-break history of the process.

As an example, consider forecasting using (8) with the α_i^*, β_i^* terms absent (the shifts are not modelled). Forecasts of y_{T+j} , $j \geq 0$ made pre-break at time $\tau - 1$ rapidly become poor as described above. The outcome is moderated a little if the forecast origin is moved forward to τ , or $\tau + 1, \tau + 2, \dots$ and the model estimates updated on the extended sample, but the impact of the post-break observations on the estimated model parameters (say, $\hat{\alpha}_i, \hat{\beta}_i$) will initially be slight. By way of contrast, consider forecasting with a model that imposes fourth-differencing, and for simplicity suppose the seasonality is not trending either pre- or post-break. Forecasts made at $\tau - 1, \tau, \dots, \tau + 3$ will go awry, but thereafter, as suggested by the pattern for $E[\Delta_4 y_t]$ set out above, the new pattern of deterministic seasonals ($\alpha_i + \alpha_i^*$) will be forecast for y_t . The model based on (8) will continue to forecast the pattern based on the α_i , subject to the small adjustments referred to above from updating.

Structural breaks in seasonals over the forecast period indicate that imposing unit roots may improve forecast accuracy even when this does not appear to be warranted on the basis of the outcomes of unit root tests.

5 Empirical analysis

We consider two examples, one for Norwegian total household real disposable income, the other for UK consumers' expenditure on non-durables and services. Both are quarterly, in constant prices, seasonally unadjusted, and we investigate their natural logarithms.

5.1 Norwegian income

This series has been analysed by Franses and Paap (1994). Time series plots of various filters of the (logarithm) of Norwegian income are shown in figure 1. The level exhibits a clear, albeit irregular or evolving, seasonal pattern. The series still appears to exhibit a seasonal pattern after regression on deterministic seasonal dummies, suggesting that the seasonality is indeed not constant over time. Taking fourth differences appears to remove much of the seasonality in the series, while the bi-annual filtered series exhibits seasonality at the beginning and end of the sample, but not over the middle range.

Another means of obtaining an indication of the amount of seasonal variation in a series is to regress the first difference (of the logarithm) of the series on a set of seasonal dummies (Miron, 1994):

$$\Delta y_t = \delta_1 D_{1,t} + \delta_2 D_{2,t} + \delta_3 D_{3,t} + \delta_4 D_{4,t} + u_t. \quad (11)$$

The $\hat{\delta}_i$ estimate the quarterly growth rate in season i : $D_{i,t}$ takes the value 1, for observations in season i , and 0 otherwise. Franses (1996) cautions against conducting statistical inference on (11), and we use this regression as a purely descriptive device. Informal comparison of the $\hat{\delta}_i$ estimated over sub-samples can be informative about the constancy of the seasonal pattern. We estimate (11) over an initial sub-sample, and then update the estimates as the sample period is successively augmented by one observation each time. The series of estimates we obtain for the $\hat{\delta}_i$'s are plotted in figure 3. It is apparent that the fourth quarter quarterly growth rates go from being negative to positive, and cross over with the second quarter rates, indicating a non-constant seasonal pattern.

Table 1 describes the results of applying the HEGY testing procedure outlined in section 2 to the Norwegian income series. The first row for the full sample and $p = 1$ confirms the findings of Franses and Paap (1994), that there appear to be zero frequency and bi-annual roots (the t_{z1} and t_{z2} test statistics are both less than the 5% critical values: significance at the 5% and 1% levels is denoted by * and ** respectively). The second row indicates evidence of serially-correlated residuals when the lagged-dependent variable is dropped, invalidating the rejection of the bi-annual root at the 1% level. All these regressions include a constant, linear trend, and deterministic seasonals.

Since we intend to use the period 1988:1–1993:2 to calculate empirical measures of forecast accuracy based on the outcomes of the HEGY testing procedure, the third and fourth rows of table 1 show the test results for the period 1967:1–1987:4 (prior to the creation of lags). For $p = 1$ we again do not reject zero frequency and bi-annual roots. Although the t -value on the lagged dependent variable is 2.528 (last column of the table), its deletion (row 4) does not lead to significant serial correlation on the LM-test for orders 1 through to 4. However, the portmanteau test will lack power for specific orders of autocorrelation, and the testing procedure is taken to indicate zero and bi-annual frequency roots. We now explore the forecast performance of models that impose unit roots at various frequencies, and in particular, whether imposing roots on the basis of the unit-root testing procedure yields significant gains.

All the forecasting models include a constant and deterministic seasonals, but no linear trend.

We then undertook a detailed analysis of the accuracy of forecasting various linear transformations of the data for models which impose unit roots at different frequencies. The linear transformations being predicted relate to various seasonal (and non-seasonal) differencing operators. For each particular model (defined by the frequency of the root(s) imposed) we calculate empirical forecast accuracy measures for

Table 1: Testing for unit roots in Norwegian income.

Sample Period	t_{z_1}	t_{z_2}	F_{z_3, z_4}	p	LM(4)	$t_{\Delta_4 y_{t-i}}, i = 1, \dots, p$
1967:1 - 1993:2	-1.360	-2.335	6.966*	1	1.026	3.057
1967:1 - 1993:2	-1.304	-3.547**	11.73**	0	2.561*	
1967:1 - 1987:4	-0.619	-2.192	7.048*	1	0.320	2.528
1967:1 - 1987:4	-0.202	-3.298*	9.877**	0	1.627	

all the linear transforms. We present empirical root mean squared forecast error (RMSFE) and generalized forecast error second moment (GFESM) calculations for horizons of l to 8-steps ahead, calculated by moving the forecast origin through the sample beginning with 1987:4, and either with, or without, updating the coefficient estimates each time.

For example, without updating, the model is estimated once and for all on the available sample (1967:1–1987:4, less periods lost through creating lags) and l to 8-step ahead forecasts are then generated (i.e. forecasts of the value of the process in 1988:1 through to 1989:4). The origin is then moved to 1988:1 and the exercise is repeated: l to 8-step ahead forecasts are again generated. This continues and dynamic forecasts are calculated up to 8-steps ahead subject to data availability for forecast error calculations. Thus, when 1992:4 is the origin we have l and 2-step forecasts (of 1993:1 and 1993:2, respectively), and when 1993:1 is the origin, only a l -step forecast of 1993:2.

With updating, the model is re-estimated on data up to and including the forecast origin before generating forecasts, but the form of the model remains unchanged.

All the models include one lagged value of the dependent variable, consistent with the HEGY testing procedure results in table 1, and a constant and three deterministic seasonals. The linear time trend was found to be insignificant and excluded. Table 1 indicates imposing a zero-frequency root and possibly a bi-annual root, though the evidence for the latter is less clear-cut.

The baseline is taken to be the unrestricted AR(5) model estimated once only over the period 1967:1–1987:4. Relative to this, updating the model each period implies small overall gains on the basis of GFESM, mainly from improvements at short horizons. However the impact of updating when accuracy is measured by MSFE is seen to depend on the linear transformation being predicted - for bi-annual changes, the consequences of updating are mixed, for seasonal and annual changes almost unambiguously bad, and for first differences yield improvements of 5% to 13% - highlighting the ambiguity that can arise from the use of such a measure, as discussed in section 3.

There are significant gains to imposing the zero-frequency root strongly suggested by the HEGY procedure (table 2). This yields large gains (up to 30% on levels and annual differences, 8-steps ahead) on all transformations other than first differences, where there is little gain at any horizon. Updating the model each period (not shown) does result in gains in the accuracy of predicting first differences, but these are still approximately only half as large as the percentage gains in predicting other linear transformations of the data.

We also find that imposing a bi-annual root has an ambiguous impact on the accuracy with which bi-annual differences are predicted, whether the model is updated or not, but that judged by the GFESM the results are an improvement on the unrestricted model overall.

Imposing seasonal roots yields the least benefit to predicting the seasonal transform when the model is not updated. There are substantial gains to predicting first differences, substantially greater than when either of the unit roots suggested by the HEGY procedure is imposed.

Imposing all roots appears to be optimal on all forecast evaluation criteria (see table 3).

A discussion of some of the features of the results for Norwegian income shared with UK consump-

Table 2: RMSFEs for model with zero frequency root imposed and no updating compared to benchmark model

j	x_j	$(1 - L) x_j$	$(1 + L) x_j$	$(1 + L^2) x_j$	$(1 - L^4) x_j$	GFESM
1	-0.682					-0.682
2	-1.594	0.095	-2.214			-1.056
3	-4.747	0.107	-4.756	-4.937		-1.640
4	-13.591	-1.289	-11.196	-8.819		-2.946
5	-13.965	-1.062	-19.060	-12.723	-17.448	-4.485
6	-15.546	-0.512	-17.429	-19.523	-27.445	-7.645
7	-18.721	-0.011	-19.715	-19.276	-30.494	-8.054
8	-30.613	-1.641	-27.333	-26.579	-32.653	-8.595

Table 3: RMSFEs for model with all roots imposed and no updating compared to benchmark model

j	x_j	$(1 - L) x_j$	$(1 + L) x_j$	$(1 + L^2) x_j$	$(1 - L^4) x_j$	GFESM
1	-12.793					-12.793
2	-11.316	-15.237	-13.767			-14.944
3	-9.307	-14.762	-11.872	-11.516		-14.903
4	-11.941	-16.570	-10.913	-13.402		-15.000
5	-30.469	-21.394	-26.549	-20.391	-20.869	-16.342
6	-32.160	-25.183	-32.063	-25.367	-34.389	-18.283
7	-29.240	-27.870	-29.391	-24.610	-34.875	-18.547
8	-34.091	-34.212	-28.874	-29.205	-36.831	-18.993

tion is postponed to section 5.3.

5.2 UK consumers' expenditure

The data here are taken from Hendry (1994) and are shown in figure 2. The estimates obtained from applying (11) to the UK data (in the same way as described above for the Norwegian income series) are recorded in figure 4. Again, a non-constant seasonal pattern is apparent: the second and fourth quarter growth rates are both clearly trending.

Table 4: Testing for unit roots in UK real consumers' non-durable expenditure.

Sample Period	t_{z_1}	t_{z_2}	F_{z_3, z_4}	p	LM(4)	$t_{\Delta_4 y_{t-i}}, i = 1, \dots, p$
1961:1 - 1992:4	-3.258	-2.496	2.923	5	0.738	3.972, 1.925, -0.969 -1.848, 4.040
1961:1 - 1992:4	-2.984	-3.655**	6.931*	2	2.683*	3.086, 1.843
1961:1 - 1986:4	-3.013	-3.564**	5.518	2	1.726	2.225, 2.024

Over the whole sample, 5 lags of the dependent variable ($p = 5$) are necessary to whiten the residuals of the HEGY regression on the portmanteau F -test. In that case, the constant, trend and Q_1 seasonal dummy are significant. With $p = 5$, we are unable to reject there being unit roots at the zero, bi-annual and seasonal frequencies, which is at odds with the sub-sample finding of unit roots at the zero and seasonal frequencies only. The table shows the sensitivity of the results to the augmentation, when $p = 2$ as in the sub-sample then we can reject the bi-annual root at the 1% level, although there is evidence of

serial correlation in the residuals at the 5% level. This set of findings suggests there is a problem of shifts in behaviour, including seasonality, which is well documented elsewhere (see Hendry, 1994).

The benchmark for comparison is the unrestricted model (constant and seasonals but excluding a trend) for which the estimates are estimated once only (no updating).

Imposing the zero-frequency unit root, as indicated by the HEGY testing procedure, yields gains in forecast accuracy of 3% to 5% at horizons of 2 to 4-steps ahead in terms of predicting the levels of the series, and the bi-annual and seasonal transformations (see table 5). Imposing the unit root and updating (not shown) improves predictions of first differences but worsens predictions of other transformations.

Table 5: RMSFEs for model with zero frequency root imposed and no updating as % of benchmark model

j	x_j	$(1 - L) x_j$	$(1 + L) x_j$	$(1 + L^2) x_j$	$(1 - L^4) x_j$	GFESM
1	-1.876					-1.946
2	-3.187	-1.351	-3.049			-1.846
3	-3.830	-1.022	-3.851	-3.640		-1.440
4	-5.018	-1.653	-4.809	-4.831		-1.326
5	-5.749	-3.041	-5.567	-5.253	-5.808	-1.305
6	-6.886	-3.313	-6.507	-6.237	-7.696	-1.446
7	-7.713	-2.565	-7.391	-6.957	-9.455	-1.396
8	-8.885	-3.306	-8.270	-7.927	-11.271	-1.353

In tune with the testing procedure, imposing seasonal roots yields gains at all but the shortest horizons, and particularly for first differences at $j = 3$ and $j = 4$ (table 6). The gains for the seasonal transform are only in the range 2% to 3% at all horizons. Updating generally worsens forecast performance relative to the benchmark, except for first differences.

Table 6: RMSFEs for model with seasonal frequency roots imposed and no updating as % of benchmark model

j	x_j	$(1 - L) x_j$	$(1 + L) x_j$	$(1 + L^2) x_j$	$(1 - L^4) x_j$	GFESM
1	1.343					0.844
2	-0.455	1.109	0.003			-0.911
3	-4.325	-8.137	-2.420	-2.611		-3.872
4	-5.398	-9.169	-4.606	-2.724		-6.979
5	-1.844	-3.361	-3.145	-2.129	-2.789	-9.688
6	-1.130	-3.632	-1.450	-2.296	-1.870	-10.319
7	-3.959	-12.447	-2.848	-3.033	-3.506	-10.079
8	-5.083	-14.481	-4.655	-3.268	-4.734	-9.668

We also find that imposing all roots is just about as good as imposing the zero and seasonal roots at the same time, as suggested by HEGY, and that in either case updating tempers the gains.

5.3 Summary

The analysis in section 4 helps explain the pattern of forecast gains and losses for both Norwegian and UK data from imposing zero frequency and bi-annual and seasonal unit roots. As a summary, at a horizon of 8-steps ahead (recording gains relative to the benchmark - unrestricted - models as positive):

The effect of updating estimates sequentially is usually small compared to the effects of imposing unit roots, so we focus on the latter. The analysis in section 4 predicts that the effect of updating is likely

Table 7: Summary table of impact of imposing roots on GFESM for 8-step ahead forecast horizons. Expressed as % gain (+) or loss (−) relative to the models with no roots imposed

Frequency of roots imposed	Norwegian income		UK consumption	
	No updating	Updating	No updating	Updating
None		3		0
Zero frequency root	9	11	1	3
Seasonal root	10	11	10	8
Bi-annual root	3	5	2	2
All roots	19	18	10	10
Unit roots suggested by testing procedure:				
Zero and bi-annual	11	12		
Zero and seasonal			10	10

to be small, if the DGP is approximately one of shifts in deterministic seasonals. As suggested by the comparison in (10), enforcing fourth differencing reduces the impact of any levels shifts in seasonality (or intercepts) to transient four quarter (or one period) ‘blips’, and reduces trend shifts to level shifts. When forecasting, such reductions in breaks can have major impacts on forecast accuracy: here, imposing all roots induces gains of approximately 10%–20% at 8-steps ahead as measured by the GFESM, relative to unrestricted models fitted in levels. For UK consumption the benefits accrue mainly from imposing the seasonal root, although the testing procedure also indicated a zero-frequency root, and possibly a bi-annual root as well (based on the whole sample). For Norwegian income forecast accuracy is enhanced by imposing roots at the zero and seasonal frequencies rather than the zero and bi-annual frequencies indicated by the testing procedure.

This suggests that there were changes in seasonality over the relevant periods, and is consistent with the evidence for changing seasonal patterns in UK consumption noted by Hendry (1994) and Harvey and Scott (1994). We interpret an important part of the changing seasonal patterns as resulting from shifts in the seasonal means so that fourth-differencing works for the reasons set out in section 4. Fourth-differencing would also be appropriate if the evolving seasonality was stochastic, emanating from roots at the bi-annual and seasonal frequencies. Generally, though, seasonal unit roots indicate that the amplitude of the seasonal cycle is unbounded in the long run and that the phase of the cycle can drift without limit, features which are at odds with the observed seasonal patterns of many series. For example Pollock (1995) notes that the phase of the cycle for consumers’ expenditure does not exhibit this sort of behaviour - the annual boom invariably occurs at Christmas time.

Finally, the dangers inherent in using any specific MSFE measure to evaluate the impact of imposing the different combinations of roots are apparent from our detailed results (not reported): no linear transform’s MSFE always had the same sign of gain as the GFESM, and at worst one transform could show a 10% gain on MSFE simultaneously with another reporting a 10% loss.

6 Conclusions

The main message of the paper is that building univariate forecasting models for quarterly, seasonally unadjusted data, conditioned on the outcome of unit-root testing procedures such as HEGY may not be useful when there are shifts in seasonal, or other aspects of, behaviour. In the empirical examples considered, imposing roots at all frequencies led to at least as good a forecasting performance as imposing the smaller number of roots suggested by the HEGY procedure (substantially better for Norwegian income, and no worse for UK consumption, see table 7). Models that imposed all roots and were estimated only

once were as good as other models that imposed a smaller number of roots and were continually updated. The theoretical analysis suggested interpreting that outcome as indicating shifts in the deterministic factors over the forecast period.

Finally, we reiterate that MSFEs are an unreliable guide to comparing forecast accuracy unless there is unique data transform of interest. For various combinations of imposed unit roots, one transform could show a gain while another showed a loss.

References

- Armstrong, J. S., and Fildes, R. A. (1995). Correspondence: On the selection of error measures for comparisons among forecasting methods. *Journal of Forecasting*, **14**, 67–72.
- Box, G. E. P., and Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.
- Christoffersen, P. F., and Diebold, F. X. (1997). Optimal prediction under asymmetric loss. *Econometric Theory*, **13**, 808–817.
- Clements, M. P., and Hendry, D. F. (1993a). On the limitations of comparing mean squared forecast errors: A reply. *Journal of Forecasting*, **12**, 669–676.
- Clements, M. P., and Hendry, D. F. (1993b). On the limitations of comparing mean squared forecast errors (with discussion). *Journal of Forecasting*, **12**, 617–637. Reprinted in T.C. Mills (ed.), *Economic Forecasting*. Edward Elgar, 1999.
- Clements, M. P., and Hendry, D. F. (1995a). Forecasting in cointegrated systems. *Journal of Applied Econometrics*, **10**, 127–146. Reprinted in T.C. Mills (ed.), *Economic Forecasting*. Edward Elgar, 1999.
- Clements, M. P., and Hendry, D. F. (1995b). A reply to Armstrong and Fildes. *Journal of Forecasting*, **14**, 73–75.
- Clements, M. P., and Hendry, D. F. (1996a). Intercept corrections and structural change. *Journal of Applied Econometrics*, **11**, 475–494.
- Clements, M. P., and Hendry, D. F. (1996b). Multi-step estimation for forecasting. *Oxford Bulletin of Economics and Statistics*, **58**, 657–684.
- Clements, M. P., and Hendry, D. F. (1998a). Forecasting economic processes. *International Journal of Forecasting*, **14**, 111–131.
- Clements, M. P., and Hendry, D. F. (1998b). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.
- Cox, D. R. (1961). Prediction by exponentially weighted moving averages and related methods. *Journal of the Royal Statistical Society, B*, **23**, 414–422.
- Ericsson, N. R. (1992). Parameter constancy, mean square forecast errors, and measuring forecast performance: An exposition, extensions, and illustration. *Journal of Policy Modeling*, **14**, 465–495.
- Fildes, R. A., and Makridakis, S. (1988). Forecasting and loss functions. *International Journal of Forecasting*, **4**, 545–550.
- Franses, P. H. (1996). Recent advances in modelling seasonality. *Journal of Economic Surveys*, **10**, 299–345.
- Franses, P. H., and Koehler, A. B. (1994). Model selection strategies for time series with increasing seasonal variation. revised version of Econometric Institute Report 9308, Erasmus University Rotterdam.

- Franses, P. H., and Paap, R. (1994). Model selection in periodic autoregressions. *Oxford Bulletin of Economics and Statistics*, **56**, 421–439.
- Franses, P. H., and Vogelsang, T. J. (1997). On seasonal cycles, unit roots and mean shifts. *Review of Economics and Statistics*. Forthcoming.
- Granger, C. W. J. (1969). Prediction with a generalized cost of error function. *Operations Research Quarterly*, **20**, 199–207.
- Granger, C. W. J. (1993). Comment on ‘On the limitations of comparing mean squared forecast errors’, by M.P. Clements and D.F. Hendry. *Journal of Forecasting*, **12**, 651–652.
- Hall, A. (1989). Testing for a unit root in the presence of moving average errors. *Biometrika*, **76**, 49–56.
- Harvey, A. C., and Scott, A. (1994). Seasonality in dynamic regression models. *Economic Journal*, **104**, 1324–1345.
- Hendry, D. F. (1994). HUS revisited. *Oxford Review of Economic Policy*, **10**, 86–106.
- Hendry, D. F., and Neale, A. J. (1991). A Monte Carlo study of the effects of structural breaks on tests for unit roots. In Hackl, P., and Westlund, A. H. (eds.), *Economic Structural Change, Analysis and Forecasting*, pp. 95–119. Berlin: Springer-Verlag.
- Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B. S. (1990). Seasonal integration and cointegration. *Journal of Econometrics*, **44**, 215–238.
- Hylleberg, S., Jørgensen, C., and Sørensen, N. K. (1993). Seasonality in macroeconomic time series. *Empirical Economics*, **18**, 321–325.
- Miron, J. A. (1994). The economics of seasonal cycles. In Sims, C. A. (ed.), *Advances in Econometrics, Sixth World Congress of the Econometric Society*. Cambridge: Cambridge University Press.
- Nelson, C. R., and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics*, **10**, 139–162.
- Nelson, H. L., and Granger, C. W. J. (1979). Experience using the Box-Cox transformation when forecasting economic time series. *Journal of Econometrics*, **10**, 57–69.
- Osborn, D. R. (1990). A survey of seasonality in UK macroeconomic variables. *International Journal of Forecasting*, **6**, 327–336.
- Paap, R., Franses, P. H., and Hoek, H. (1997). Mean shifts, unit roots and forecasting seasonal time series. *International Journal of Forecasting*, **13**, 357–368.
- Perron, P. (1989). The Great Crash, the oil price shock and the unit root hypothesis. *Econometrica*, **57**, 1361–1401.
- Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, **8**, 153–162.
- Pollock, D. S. G. (1995). Metaphors for time-series analysis. *International Journal for Mathematical Education in Science and Technology*, **26**, 501–515.
- Smith, J., and Otero, J. (1997). Structural breaks and seasonal integration. *Economic Letters*. Forthcoming.
- Stoica, P., and Nehorai, A. (1989). On multistep prediction error methods for time series models. *Journal of Forecasting*, **8**, 357–368.
- Tiao, G. C., and Xu, D. (1993). Robustness of maximum likelihood estimates for multi-step predictions: The exponential smoothing case. *Biometrika*, **80**, 623–641.
- Tsay, R. S. (1993). Comment: Adaptive forecasting. *Journal of Business and Economic Statistics*, **11**, 140–142.
- Weiss, A. A. (1991). Multi-step estimation and forecasting in dynamic models. *Journal of Econometrics*,

48, 135–149.

Weiss, A. A. (1996). Estimating time series models using the relevant cost function. *Journal of Applied Econometrics*, **11**, 539–560.

Weiss, A. A., and Andersen, A. P. (1984). Estimating time series models using the relevant forecast evaluation criterion. *Journal of the Royal Statistical Society, A*, **147**, 484–487.

Zellner, A. (1986). A tale of forecasting 1001 series: The Bayesian Knight strikes again. *International Journal of Forecasting*, **2**, 491–494.

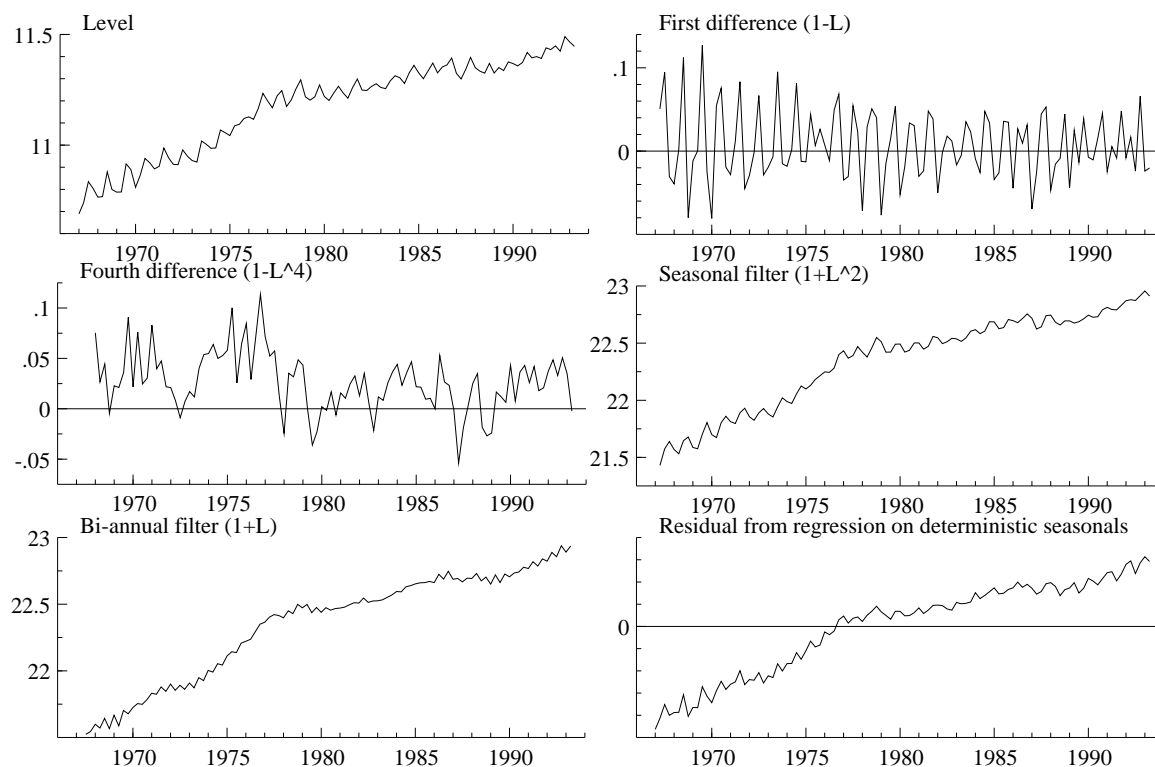


Figure 1: Total Household Real Disposable Income in Norway: The results of applying various filters

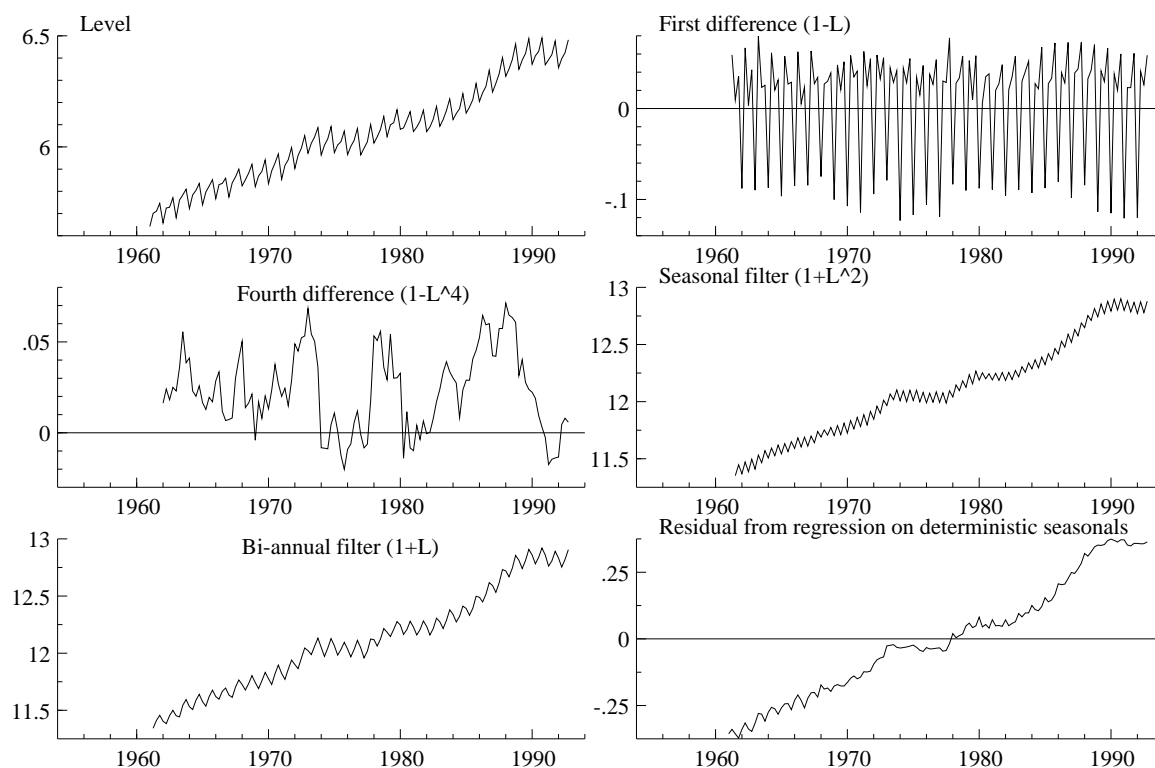


Figure 2: Real Non-durable Consumers' Expenditure in the UK: The results of applying various filters

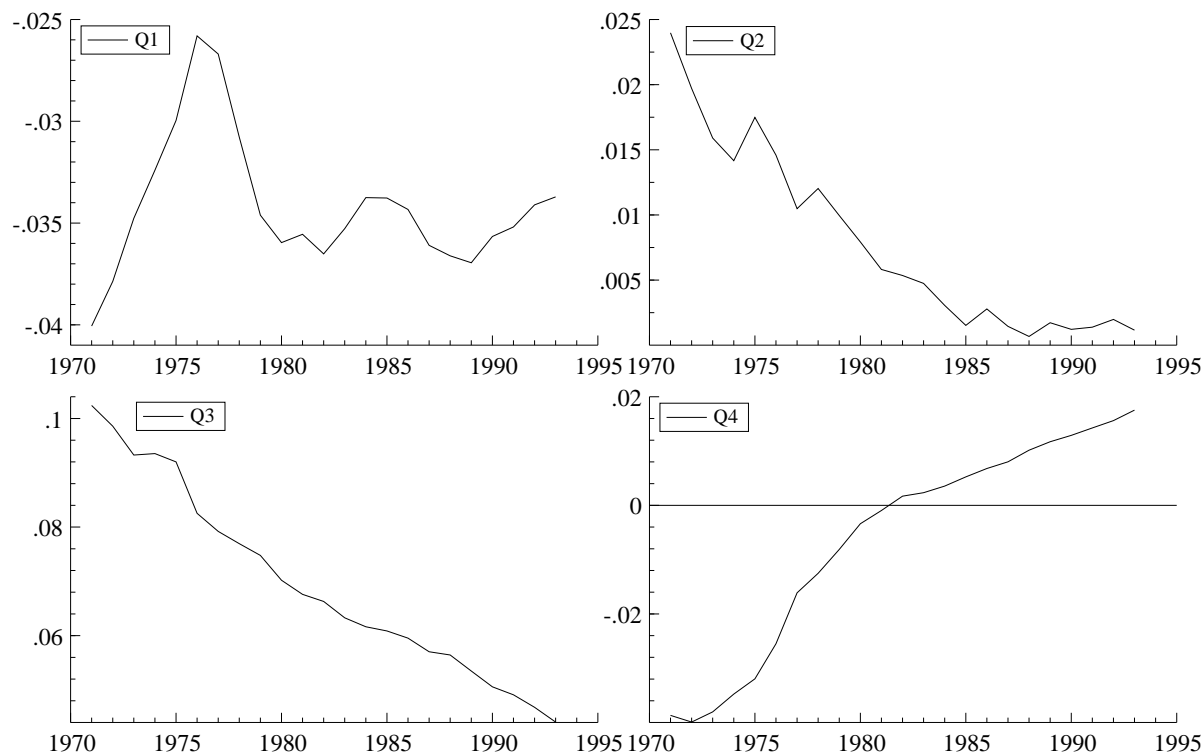


Figure 3: Estimates of quarterly growth rates of Norwegian income by season

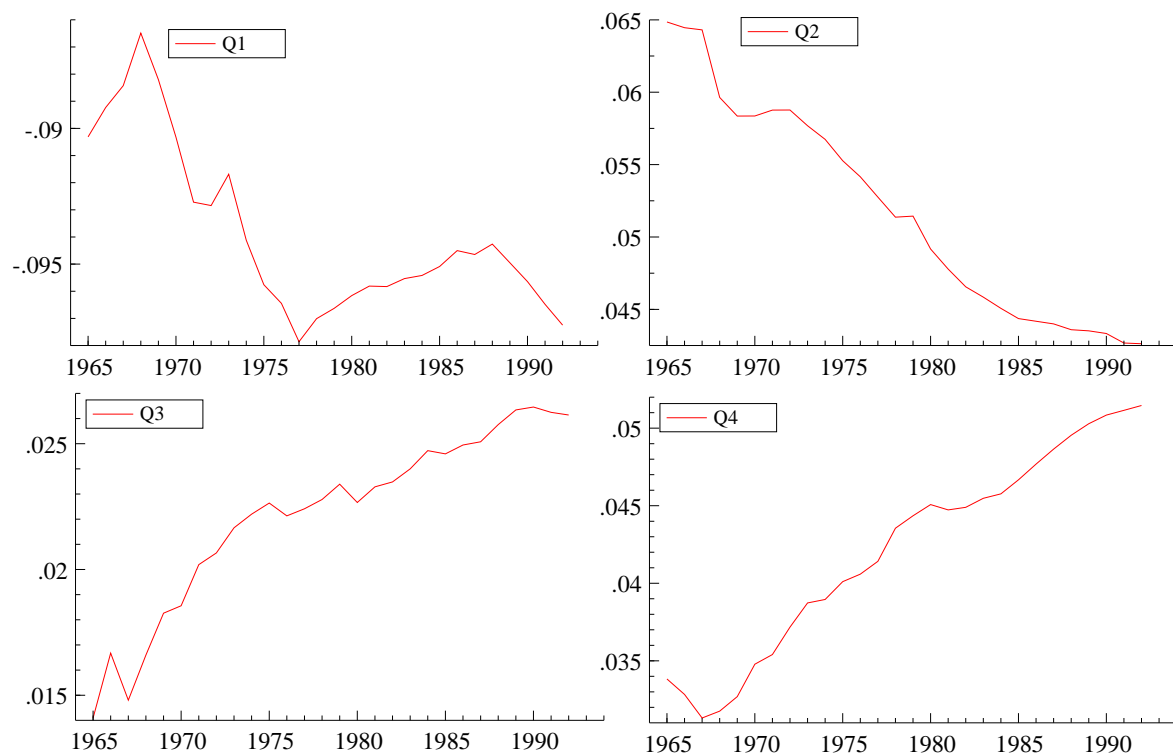


Figure 4: Estimates of quarterly growth rates of UK consumption by season