



ELSEVIER

International Journal of Forecasting 13 (1997) 357–368

*international journal
of forecasting*

Mean shifts, unit roots and forecasting seasonal time series

Richard Paap^{a,*}, Philip Hans Franses^b, Henk Hoek^c

^a*Tinbergen Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands*

^b*Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands*

^c*Tinbergen Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands*

Abstract

Examples of descriptive models for changing seasonal patterns in economic time series are autoregressive models with seasonal unit roots or with deterministic seasonal mean shifts. In this paper we show through a forecasting comparison for three macroeconomic time series (for which tests indicate the presence of seasonal unit roots) that allowing for possible seasonal mean shifts can improve forecast performance. Next, by means of simulation we demonstrate the impact of imposing an incorrect model on forecasting. We find that an inappropriate decision can deteriorate forecasting performance dramatically in both directions, and hence we recommend the practitioner to take account of seasonal mean shifts when testing for seasonal unit roots. © 1997 Elsevier Science B.V.

Keywords: Seasonal unit roots; Seasonal mean shifts; Forecasting comparison

1. Introduction

Quarterly growth rates of many macroeconomic variables observed for a few decades typically do not display a constant seasonal pattern over time, see e.g. Hylleberg (1994) for some illustrative graphs. Several time series models have been proposed to describe such unstable patterns. One approach is given by imposing unit roots at the seasonal frequencies in the autoregressive [AR] part of a model, which corresponds to the assumption of a changing seasonal pattern caused by seasonal stochastic trends, see Hylleberg et al. (1990). Another approach is to introduce periodically varying AR parameters, see among others Osborn (1988) and Franses (1996). Yet another method amounts to allowing deterministic changes in the seasonal means. Such deterministic

seasonal mean shifts can occur for example when two sources of data are combined or when a statistical agency introduces new measurement methods. In this paper, we focus on the first and third approach, and we compare their relative merits in terms of out-of-sample forecasting.

The influence of seasonal mean shifts on testing for seasonal unit roots has been analyzed in Ghysels (1994); Franses and Vogelsang (1997) and Franses et al. (1997) among others. It appears that neglected deterministic seasonal mean shifts bias the outcomes of seasonal unit root tests towards non-rejection, which is similar to the findings in Perron (1990) for the nonseasonal case. It seems therefore necessary to allow for possible seasonal mean shifts when testing for seasonal unit roots in order to avoid overdifferencing. Since overdifferencing usually results in many lagged regressors in AR models, one may expect forecasting performance to depend on the

*Corresponding author. e-mail: paap@tir.few.eur.nl

appropriate description of changing seasonality. In this paper we investigate through simulations the effects of neglecting and imposing seasonal mean shifts on the forecasting performance of univariate AR models. We compare forecasts from AR models with one or more seasonal unit roots with forecasts from models with deterministic seasonal mean shifts.

The outline of our paper is as follows. In Section 2 we discuss the two models of interest and compare their forecast performance for three macroeconomic consumption series to illustrate the focus of this paper. To show the effects of seasonal mean shifts on the forecasting performance of AR models with seasonal unit roots and vice versa, we perform several Monte Carlo experiments in Section 3. In Section 4, we conclude this paper by giving some practical suggestions.

2. Seasonal unit roots and seasonal mean shifts

A procedure to test for seasonal unit roots in a quarterly time series, $y_t, t = 1, \dots, T$, is developed by Hylleberg et al. (1990) [HEGY]. This procedure is based on the auxiliary regression model

$$\begin{aligned} \Delta_4 y_t = & \sum_{s=1}^4 \mu_s D_{st} + \gamma T_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} \\ & + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \sum_{i=1}^k \phi_i \Delta_4 y_{t-i} + \varepsilon_t, \end{aligned} \tag{1}$$

where D_{st} represent the usual seasonal dummies ($s=1, 2, 3, 4$), T_t is a deterministic trend term ($T_t=0, 1, 2, \dots$), $\Delta_i y_t = y_t - y_{t-i}$, and where ε_t is assumed to be a standard white noise process. The terms

$$\begin{aligned} y_{1,t} &= (1 + B + B^2 + B^3)y_t = y_t + y_{t-1} + y_{t-2} + y_{t-3} \\ y_{2,t} &= (-1 + B - B^2 + B^3)y_t \\ &= -y_t + y_{t-1} - y_{t-2} + y_{t-3} \\ y_{3,t} &= (-1 + B^2)y_t = -y_t + y_{t-2}, \end{aligned} \tag{2}$$

with $B^i y_t = y_{t-i}$, can be seen as error correction terms such that the insignificance of the π_i parameters corresponds to the presence of seasonal and

nonseasonal unit roots. In fact, if $\pi_1 = 0$, the series has a unit root at the zero frequency, which implies that the series has a nonseasonal stochastic trend. The appropriate differencing filter in that case is $(1 - B)$. A unit root at the bi-annual frequency (-1) , i.e. two cycles per year, corresponds with $\pi_2 = 0$. The presence of this unit root implies that the differencing filter $(1 + B)$ is appropriate. If $\pi_3 = 0$ and $\pi_4 = 0$ the series contains the roots i and $-i$, i.e. seasonal unit roots at the annual frequencies, which corresponds to one cycle per year and to the filter $(1 + B^2)$. Applying OLS to (1), the significance of the π_i parameters can be checked by t - and F -tests. Simulated critical values of t -tests for the significance of the π_i 's and an F -test for the significance of π_3 and π_4 are tabulated in Hylleberg et al. (1990).

A typical outcome of an application of the HEGY method is that macroeconomic time series tend to have one or more seasonal unit roots. For example, Osborn (1990) detects seasonal unit roots in six out of thirty UK macroeconomic variables. Otto and Wirjanto (1990) obtain similar results for Canadian macroeconomic time series and Hylleberg et al. (1993) find that several country-specific Gross Domestic Product series have one or more seasonal unit roots.

The influence of seasonal mean shifts on seasonal unit root inference is analyzed in several recent papers. For example, Franses and Vogelsang (1997) find that the evidence for the bi-annual unit root -1 in USA Industrial Production disappears, when a seasonal mean shift is added to the test regression (1). With Bayesian methods, Franses et al. (1997) show that the evidence for seasonal unit roots in three consumption series tends to disappear when one allows for a possible seasonal mean shift. The series in this last paper are the log of real Total Consumption of Sweden, 1963.1–1988.4, the log of real Total Consumption of the UK, 1955.1–1988.4, and the log of real Nondurable Consumption in the USA, 1947.1–1991.4. For expository purposes, we re-analyze these series within the light of out-of-sample forecasting. Table 1 shows the results of the standard HEGY test for unit roots using test regression (1). The test regression contains seasonal dummies and a trend and the order k is based on the Lagrange Multiplier [LM] test for first-to-fourth order serial correlation in the residuals. From the first

Table 1

HEGY test results and outcomes of the tests for the most likely breakpoint in a first difference model with seasonal mean shifts for three consumption series

Country	Period	HEGY test ^a				Breakpoint test ^b	
		<i>k</i>	<i>t</i> (π_1)	<i>t</i> (π_2)	<i>F</i> (π_3, π_4)	<i>p</i>	estimated τ
Sweden	63.1-88.4	8	-1.54	-1.25	3.64	5	1980.3
UK	55.1-88.4	8	-1.69	-1.14	2.81	9	1977.2
USA	47.1-91.4	6	-0.43	-2.24	5.06	9	1956.3

^a*k* denotes the order of the model in test Eq. (1). The 5% critical values of *t*(π_1), *t*(π_2), and *F*(π_3, π_4) are -3.53, -2.94, and 6.60 respectively.

^b*p* denotes the order of the model in Eq. (5).

part of Table 1 we note that we cannot reject the presence of the nonseasonal and the three seasonal unit roots for all three series, which indicates that we have to transform the series with the Δ_4 filter. This conclusion does not change if insignificant lags of fourth differences are removed in the *k*th order HEGY test regression. The AR orders of the models in fourth differences turn out to be 8 for Sweden, 8 for the UK, and 6 for the USA.

Franses et al. (1997) consider an alternative model to describe these three consumption series, which is capable of describing seasonal mean shifts, i.e.,

$$\Delta_1 y_t = \sum_{s=1}^4 \delta_s D_{st} + \sum_{s=1}^4 \delta_s^* D_{st} [I_{t \geq \tau}] + \eta_t, \tag{3}$$

where $[I_{t \geq \tau}]$ is 1 if $t \geq \tau$ and zero otherwise, and where η_t is some AR type error process. Hence, the quarterly intercepts are δ_s for $t < \tau$ and $\delta_s + \delta_s^*$ for $t \geq \tau$, $s = 1, 2, 3, 4$. Note that if $\sum_{s=1}^4 \delta_s^* \neq 0$, the annual drift term changes from $\sum_{s=1}^4 \delta_s$ to $\sum_{s=1}^4 (\delta_s + \delta_s^*)$. Model (3) allows for a one-time change at time τ in the seasonal means. The higher order dynamics enter the model by an AR(*p*) structure on η_t in (3), or

$$\eta_t = \sum_{i=1}^p \phi_i \eta_{t-i} + \varepsilon_t. \tag{4}$$

Combining (3) with (4) leads to the following model for $\Delta_1 y_t$,

$$\Delta_1 y_t = \sum_{s=1}^4 \mu_s D_{st} + \sum_{s=1}^4 \mu_s^* D_{st} [I_{t \geq \tau}] + \sum_{i=1}^p (\phi_i \Delta_1 y_{t-i} + \alpha_i [I_{t=\tau-1+i}]) + \varepsilon_t, \tag{5}$$

where $[I_{t=\tau-1+i}]$ is 1 if $t = \tau - 1 + i$ and zero otherwise. The parameters μ_s, μ_s^* and α_i are functions of the parameters δ_s, δ_s^* , and $\phi_i, s = 1, 2, 3, 4$ and $i = 1, \dots, p$. The α_i parameters accommodate the dynamic adaptation of the changing seasonal means.

To construct the model with the seasonal mean shifts for our three consumption series, we first build a model in first differences. AR models of order $p = 5$ for Sweden, $p = 9$ for the UK, and $p = 9$ for the USA result in models without first-to-fourth order serial correlation in the residuals. In these models we add seasonal mean shifts like in (5) for different values of τ and test for the significance of these seasonal mean shifts using an *F*-test on $\mu_s^* = 0, s = 1, 2, 3, 4$ and $\alpha_i = 0, i = 1, \dots, p$. The value of τ for which the *F*-test takes the largest value is taken as an estimate for the timing of the seasonal mean shift, see Andrews (1993). The estimated values of τ , see the second part of Table 1, correspond reasonably well with the posterior modes of the breakpoint parameter in the Bayesian analysis of Franses et al. (1997), which are given by 1979.3 for Sweden, 1978.2 for the UK and 1957.4 for the USA, respectively. After imposing the seasonal mean shift we adjust the order of the AR model in first differences. The definitive order of the AR model with seasonal mean shifts is 3 for the Swedish series, 1 for the UK series and 3 for the USA series, and we use these models for out-of-sample forecasting.

To evaluate the fourth order differenced models and the seasonal mean shift models, we analyze their relative forecast performance. The sample is split up in an estimation and a validation period. For the USA and the UK we remove the last 24 observations to evaluate the forecasts. Since the seasonal mean shift for the consumption series of Sweden occurs at

Table 2

Forecasting performance of the seasonal unit root AR model versus the model with seasonal mean shifts for the three consumption series^a

Series		One-step ahead		Four-steps ahead	
	<i>period</i>	85.1–86.4	85.1–88.4	85.4–86.4	85.4–88.4
Sweden	GFESM ^b	0.48	0.61	2.59	0.52
	$F_{SUR,SS}$	0.74 (0.42)	0.02 (0.89)	1.62 (0.27)	2.47 (0.14)
	$F_{SS,SUR}$	9.01 (0.02)	9.64 (0.01)	0.36 (0.58)	0.10 (0.75)
	<i>period</i>	83.1–85.4	83.1–88.4	83.4–85.4	83.4–88.4
UK	GFESM ^b	1.22	1.17	3.96	0.80
	$F_{SUR,SS}$	3.45 (0.09)	4.44 (0.05)	4.53 (0.06)	0.91 (0.35)
	$F_{SS,SUR}$	0.87 (0.37)	0.46 (0.50)	0.00 (0.97)	1.96 (0.18)
	<i>period</i>	86.1–88.4	86.1–91.4	86.4–88.4	86.4–91.4
USA	GFESM ^b	1.40	1.24	3.23	4.77
	$F_{SUR,SS}$	4.83 (0.05)	5.74 (0.03)	6.08 (0.04)	2.32 (0.14)
	$F_{SS,SUR}$	0.30 (0.59)	0.14 (0.71)	0.12 (0.74)	0.95 (0.34)

^aThe models are estimated using the sample 63.1–84.4 for Sweden, 55.1–82.4 for the UK and 47.1–85.4 for the USA series.

^bThe top cells denote the ratio of the GFESM of the forecasts of the fourth difference model and the model with seasonal mean shifts. If the ratio exceeds 1, the model with seasonal mean shifts is preferred. $F_{SUR,SS}$ and $F_{SS,SUR}$ denote the forecast encompassing tests in Eq. (6) with *p*-values between brackets.

the end of the series, only the last 16 observations are removed. Table 2 shows the forecast performance of both models according to the Generalised Forecast Error Second Moment [GFESM] criterion. This criterion is identical to the determinant of the covariance matrix of the 1 through *h*-step ahead forecast errors. For one-step ahead forecasts this criterion simplifies to the mean squared forecast error. An advantage of this criterion is that it is invariant to non-singular, scale-preserving linear transformations of the linear forecasting model, see Clements and Hendry (1993). The top cells in Table 2 denote the ratio of the GFESM of the fourth order difference model and the seasonal mean shift model. If this ratio exceeds 1, the model with imposed mean shifts is preferred.

For the Swedish series the one-step ahead forecasts of the fourth order difference model are better than those of the seasonal mean shift model. For the UK and the USA, however, the one-step ahead forecasts of the seasonal mean shift model are better. For the USA series the same is true for the four-step ahead forecasts but for the UK and the Swedish series the model in fourth differences has a smaller GFESM when we evaluate four-step ahead forecasts for the whole forecasting period.

Apart from comparing forecast criteria, we also compare the forecasting performance of competing models using forecast encompassing tests. Let \tilde{y}_{T+h} and \tilde{y}_{T+h} denote *h*-step ahead forecasts of the seasonal unit root model and the model with a seasonal mean shift, respectively. A forecast encompassing test can be performed in the following test regression,

$$y_{T+h} - \tilde{y}_{T+h} = \beta(\tilde{y}_{T+h} - \tilde{y}_{T+h}) + u_t, \tag{6}$$

where y_{T+h} is the ‘true’ value of the series, see e.g. Clements and Hendry (1993). If β does not significantly differ from zero, the forecasts of the seasonal unit root model encompass the forecasts of the seasonal mean shift model. We denote the *F*-test for the significance of β by $F_{SUR,SS}$. Likewise, we can test whether the seasonal mean shift model encompasses the seasonal unit root model using an $F_{SS,SUR}$ -test.

Note that several situations can occur. If both the $F_{SUR,SS}$ -statistic and the $F_{SS,SUR}$ -statistic are significant, neither of the two models produces superior forecasts. The latter occurs, when only one of the *F*-test values implies rejection of the null hypothesis. Finally, the forecasts of both models are not sig-

nificantly different, if neither of the F -test values indicates rejection.

Table 2 shows the results of the encompassing tests. For the Swedish series the one-step ahead forecasts of the fourth difference model encompass the forecasts from the seasonal mean shift model. For the UK and USA series, we arrive at the opposite conclusion. For the four-step ahead forecasts, there is some evidence that the seasonal mean shift model encompasses the fourth difference model for the UK and the USA series for the period 1983.4–1985.4 and 1986.4–1988.4, respectively.

In summary, we note that although the HEGY procedure selects a model with three seasonal unit roots, the seasonal mean shift model seems to produce more adequate forecasts for two out of three series under consideration. In general, we can say that model (5) describes changing seasonality as a deterministic change in the seasonal means of a time series. This is in contrast to models with seasonal unit roots, which assume that the seasonal pattern changes due to a seasonal stochastic trend. Obviously, it is not surprising that it is often difficult to decide between the two models in practice, although the above results show that an inappropriate choice may lead to biased forecasts. Incorrectly imposing seasonal unit roots leads to overdifferencing and hence to seasonal unit roots in the moving average part of the model. This may have a negative influence on forecasting accuracy. Neglecting a seasonal unit root and imposing a deterministic seasonal mean shift assumes that forecasts of the observations in particular seasons are constant while, in fact, they are likely to change in the future. Furthermore, deterministic changes in the seasonal means can also occur outside the estimation sample and cannot be forecasted with a model like (5). To detect seasonal mean shifts in the forecasting period one could think of a recursive forecasting method in combination with Hendry's predictive failure tests or Chow tests; see Hendry (1979). Notice that in this case one assumes the availability of a hold-out sample.

In the next section we investigate the effects of neglecting seasonal mean shifts or neglecting seasonal unit roots on the forecasting performance of AR models. Additionally, we consider the forecast performance of the seasonal mean shift model and the

model with seasonal unit roots for series with recurrent seasonal mean shifts and with seasonal means which evolve as martingales. The investigation is restricted to detecting seasonal mean shifts in the sample and not in the forecast period. We decide to focus on AR models, since these are typically used in the context of seasonal unit roots and seasonal mean shifts. Furthermore, we focus on point forecasts instead of interval forecasts.

3. Monte Carlo experiments

In this section, we use Monte Carlo techniques to analyze the impact of seasonal mean shifts on forecasting performance. Several data generating processes [DGPs] are considered.

The simulation experiment is performed in the following way. First we consider for the simulated series a model with possible seasonal unit roots. We test for seasonal unit roots using Eq. (1) with $\gamma=0$. The number of lags k is determined using an LM test on first-to-fourth order autocorrelation in the residuals using a 5% significance level. We impose the unit roots indicated by the HEGY procedure, where we again adopt a 5% significance level, and we construct an AR model for the resulting transformed time series using the same LM test. The nonseasonal unit root is always imposed, since our main focus is on modeling the seasonal component.

The second model is a model in first differences with four seasonal dummies. We determine the number of lagged first differences using an LM test on first-to-fourth order autocorrelation in the residuals. In this model we include for all possible values of τ a seasonal mean shift as in (5). The estimate of the timing of the seasonal mean shift is determined by the largest value of an F -test for $\mu_s^*=0$, $s=1, 2, 3, 4$ and $\alpha_i=0$, $i=1, \dots, p$. After imposing the most likely seasonal mean shift, we construct an AR model and decide on its order again using the same LM test for serial correlation. We feel that this experiment mimics a useful empirical strategy in practice.

Finally, we generate forecasts for $\Delta_1 y_t$ from the AR model selected by the HEGY procedure and from the AR model with seasonal mean shifts. Notice that the HEGY test may also indicate that no

seasonal unit roots are needed. We consider one-step, four-step and eight-step ahead forecasts for 16, 32, and 64 quarters. The forecasts are evaluated with the GFESM criterion and with the two encompassing tests using a 5% level of significance.

The Monte Carlo experiment is split up into three parts. In Section 3.1, we consider DGPs with seasonal mean shifts. In Section 3.2, DGPs with seasonal unit roots are considered. Section 3.3 deals with other DGPs which incorporate recurrent changing seasonal means.

3.1. DGPs with seasonal mean shifts

The DGPs in this section are of the form

$$\Delta_1 y_t = \sum_{s=1}^4 \mu_s D_{st} + \sum_{s=1}^4 \mu_s^* D_{st} [I_{t \geq \tau}] + \varepsilon_t, \quad (7)$$

$$\varepsilon_t \sim NID(0, 1).$$

Since we only focus on seasonal mean shifts, we impose that $\sum_{s=1}^4 \mu_s = 0$, and $\sum_{s=1}^4 \mu_s^* = 0$, i.e., there are no annual drifts. We set the parameters as $\mu_s = \{1, -1, 1, -1\}$ and $\mu_s^* = w \mu_s$ with $w = -3, \dots, 0, \dots, 3$. The sample size $T = 136$ and the mean shift occurs at $\tau = 69$. The number of replications is 5000.

In Tables 3–6 we display results for DGPs with seasonal mean shifts. First, we consider the effect of the size of the seasonal mean shift, denoted by w , on the forecast performance in Table 3. Each cell in Table 3 denotes the percentage that an AR model selected by the HEGY procedure forecasts better than a model with seasonal mean shifts (5) according to the GFESM criterion. Several conclusions can be drawn from Table 3. If there is no seasonal mean shift (i.e., $w = 0$) the model selected by tests for seasonal unit roots forecasts better than the seasonal mean shift model in about 75% of the cases. Similar results are found for one-, four-, and eight-step ahead forecasts. For the DGPs with a seasonal mean shift, i.e. the case with $w \neq 0$, the results are different. Already for a small mean shift the seasonal mean shift model is better than the model selected by the HEGY procedure in about 80% of the cases. This percentage increases to over 90% if we evaluate more forecasts.

Similar results are obtained for the encompassing tests. In about 90% of the cases the forecasts of the

seasonal mean shift model encompasses the model with possible seasonal units. However, if the number of forecasts is small, the model with possible seasonal unit roots encompasses the seasonal mean shift model in more than 50% of the cases. Note that in case of no seasonal mean shift ($w = 0$) this percentage is about 90%, as expected. The percentage that the null hypothesis for both encompassing tests cannot be rejected is relatively high. This means that the encompassing tests do not always provide unambiguous information.

Finally, the relative performance of the models appears symmetric around $w = 0$, which suggests that the direction of the shift, i.e. positive or negative, does not seem important for forecasting.

Next, we consider the influence of the location of the seasonal mean shift within the range of observations on forecasting. Table 4 shows the outcomes of the simulation exercise for a seasonal mean shift at three different time points, 1/4, 1/2, and 3/4 of the sample size. For the one-step ahead forecasts it seems that when the timing of a seasonal mean shift is later in the sample, the forecasting performance of the seasonal mean shift model decreases slightly. A possible reason can be that the number of observations to estimate the new seasonal mean is too small to produce relatively good forecasts. The same conclusion can be drawn for the performance of the encompassing tests. For the four- and eight-step ahead forecasts we see that the model selected by the HEGY procedure performs relatively poor when the seasonal mean shift occurs in the middle of the sample, compared to shifts in the beginning or at the end of the series.

The effects of the number of seasonal means that are changing on the forecasting performance is shown in Table 5. Note that due to the restriction $\sum_{s=1}^4 \mu_s^* = 0$ it is not possible to analyze the effects of only a single change in the mean. The results indicate that the more seasons change, the better the model selected by the HEGY procedure performs, although the model with deterministic shifts outperforms the seasonal unit root model in more than 85% of the cases. Also, the percentage that forecasts of the model with possible seasonal unit roots encompass the forecasts of the seasonal mean shift model increases, when the number of changing seasonal means increases.

Table 3
The effects of neglecting seasonal mean shifts on the forecasting performance

w	Criterion ^a	One-step ahead number of periods			Four-steps ahead number of periods			Eight-steps ahead number of periods		
		16	32	64	16	32	64	16	32	64
		–3	GFESM	15.5	8.9	4.0	15.8	9.3	4.4	15.6
	$F_{SUR,SS}$	58.5	19.8	1.8	54.0	19.7	2.3	52.6	18.2	2.1
	$F_{SS,SUR}$	95.0	92.7	89.5	93.8	90.1	86.6	93.0	90.5	85.3
–2	GFESM	16.6	10.5	5.1	16.6	10.0	4.7	15.8	8.6	4.3
	$F_{SUR,SS}$	62.7	25.8	3.5	56.4	23.6	3.7	55.0	25.5	5.0
	$F_{SS,SUR}$	95.0	93.1	90.3	94.2	92.2	88.4	93.4	90.7	86.0
–1	GFESM	21.2	15.4	10.2	22.3	15.7	10.9	21.8	15.0	10.4
	$F_{SUR,SS}$	69.1	39.7	13.2	66.7	38.0	13.1	64.2	37.6	13.2
	$F_{SS,SUR}$	94.6	91.8	86.7	93.5	90.3	84.4	92.9	89.4	83.7
0	GFESM	69.6	73.9	79.5	69.4	73.5	79.7	66.1	71.4	77.8
	$F_{SUR,SS}$	94.0	92.0	89.6	94.2	92.4	90.4	93.7	92.5	90.4
	$F_{SS,SUR}$	76.6	61.3	46.0	76.5	62.0	45.9	77.0	61.6	45.6
1	GFESM	22.3	16.0	10.5	22.7	16.6	10.5	21.6	15.7	10.4
	$F_{SUR,SS}$	70.0	41.8	13.4	66.3	40.2	13.6	65.0	38.6	13.6
	$F_{SS,SUR}$	94.0	91.0	86.6	93.4	89.4	84.2	92.8	89.0	83.6
2	GFESM	16.4	10.5	5.4	16.1	10.0	5.2	15.2	8.8	4.4
	$F_{SUR,SS}$	62.5	25.2	4.0	56.7	24.9	3.9	53.7	26.0	4.9
	$F_{SS,SUR}$	94.9	93.4	89.6	94.7	92.1	88.3	93.8	91.7	86.8
3	GFESM	15.1	8.9	4.2	15.7	9.1	4.6	14.5	7.7	3.8
	$F_{SUR,SS}$	58.7	19.6	1.9	53.3	20.4	1.9	51.0	18.9	1.9
	$F_{SS,SUR}$	94.9	93.0	88.9	94.7	91.6	86.7	94.0	91.7	86.1

^aThe top cell denotes the percentage that the GFESM of the model selected by the HEGY procedure is smaller than the GFESM of the seasonal mean shift model. The middle cell denotes the percentage that the forecasts generated by the model selected by HEGY encompass the forecasts generated by the seasonal mean shift model. The bottom cell denotes the percentage that the forecasts generated by the seasonal mean shift model encompass the forecasts generated by the model selected by the HEGY procedure.

Total number of observations $T=136$.

Finally, we analyze forecasting performance across variation in the seasons. From Table 6 we notice that if we change two seasons, it does not matter which two seasons we change. Unreported simulation results show that the same conclusion can be drawn if three instead of two seasons change.

3.2. DGPs with seasonal unit roots

In the second part we analyze the relative forecast performance for DGPs containing one or more seasonal unit roots. The following DGPs are considered

$$\Phi(B)y_t = \varepsilon_t, \varepsilon_t \sim NID(0, 1) \quad (8)$$

where $\Phi(B)$ is either $(1-B^4)$, $(1+B^2)(1-B)$, $(1+B)(1-B)$ or $(1-B)$. This corresponds to three, two, one, and no seasonal unit roots, respectively.

Table 7 shows the results for the above mentioned DGPs. As a benchmark we also report the results of a DGP having no seasonal unit roots and no seasonal mean shifts. The results indicate that the models selected by the HEGY procedure produce better one- and four-step ahead forecasts. Remarkably, in the case that the DGP has seasonal unit roots, the eight-step ahead forecasts from the model selected by the

Table 4
The effects of the location of the mean shift on the forecasting performance^a

τ	Criterion ^b	One-step ahead number of periods			Four-steps ahead number of periods			Eight-steps ahead number of periods		
		16	32	64	16	32	64	16	32	64
35	GFESM	16.1	9.3	4.2	15.6	9.7	4.5	16.0	8.5	3.8
	$F_{SUR,SS}$	65.3	30.6	5.7	68.8	42.2	15.3	69.1	43.6	17.1
	$F_{SS,SUR}$	95.7	94.5	93.1	95.4	94.4	92.8	94.3	93.2	91.2
69	GFESM	16.6	10.5	5.1	16.6	10.0	4.8	15.8	8.6	4.3
	$F_{SUR,SS}$	62.7	25.8	3.5	56.4	23.6	3.7	55.0	25.8	5.0
	$F_{SS,SUR}$	95.0	93.1	90.3	94.2	92.2	88.4	93.4	90.7	86.0
103	GFESM	24.4	19.1	15.2	24.0	19.4	15.1	23.8	19.1	15.2
	$F_{SUR,SS}$	62.8	31.6	8.9	46.8	24.8	6.4	44.4	25.6	8.0
	$F_{SS,SUR}$	90.9	84.3	75.0	89.6	81.3	71.1	88.2	79.5	67.8

Total number of observations $T=136$.

^aThe DGP is given by Eq. (7).

^bSee footnote ^a of Table 3.

Table 5
The effects of the number of quarters that are changing on the forecasting performance^a

Number of quarters	Criterion ^b	One-step ahead number of periods			Four-steps ahead number of periods			Eight-steps ahead number of periods		
		16	32	64	16	32	64	16	32	64
2	GFESM	12.5	6.0	2.3	12.1	5.9	2.3	11.9	5.3	2.0
	$F_{SUR,SS}$	44.5	12.0	1.1	45.0	12.8	1.1	41.2	11.4	1.1
	$F_{SS,SUR}$	94.0	92.1	88.9	94.2	92.2	88.0	93.9	91.3	86.5
3	GFESM	15.1	8.3	4.0	14.8	8.2	4.4	15.7	8.1	4.1
	$F_{SUR,SS}$	50.9	18.8	2.6	51.6	19.8	3.1	49.2	18.6	3.0
	$F_{SS,SUR}$	93.9	91.6	87.4	93.8	91.4	87.1	94.1	91.4	86.4
4	GFESM	16.6	10.5	5.1	16.6	10.0	4.8	15.8	8.6	4.3
	$F_{SUR,SS}$	62.7	25.8	3.5	56.4	23.6	3.7	55.0	25.8	5.0
	$F_{SS,SUR}$	95.0	93.1	90.3	94.2	92.2	88.4	93.4	90.7	86.0

Total number of observations $T=136$.

^aThe DGP is given in (7) with $\mu_s^* = \{-2, 2, 0, 0\}$, $\mu_s^* = \{-1, 2, -1, 0\}$, and $\mu_s^* = \{-2, 2, -2, 2\}$ for 2, 3, and 4 changing quarters respectively.

^bSee footnote ^a of Table 3.

HEGY procedure are in more than 50% of the cases worse than forecasts from the model with the seasonal mean shift. From the outcomes of the second and third DGP of the table it seems that neglecting the root -1 is worse than neglecting the roots i and $-i$, if one is interested in one- and

four-step ahead forecasts. In about 90% of the cases the forecasts of the model with possible seasonal unit roots encompass the forecasts of the seasonal mean shift model. The percentage that the seasonal mean shift model encompasses the model with possible seasonal unit roots is larger for DGPs with only one

Table 6
The effects of the choice of quarters we allow to change on forecasting performance^a

Changing quarters	Criterion ^b	One-step ahead number of periods			Four-steps ahead number of periods			Eight-steps ahead number of periods		
		16	32	64	16	32	64	16	32	64
(1,2)	GFESM	12.5	6.0	2.3	12.1	5.9	2.3	11.9	5.3	2.0
	$F_{SUR,SS}$	44.5	12.0	1.1	45.0	12.8	1.1	41.2	11.4	1.1
	$F_{SS,SUR}$	94.0	92.1	88.9	94.2	92.2	88.0	93.9	91.3	86.5
(1,4)	GFESM	11.8	6.4	2.4	11.5	6.0	2.4	11.9	5.8	2.3
	$F_{SUR,SS}$	44.5	12.2	1.2	45.0	13.7	1.3	42.2	11.8	1.4
	$F_{SS,SUR}$	94.2	92.6	88.5	94.2	92.2	87.6	93.5	90.4	85.2
(2,3)	GFESM	12.0	6.1	2.5	11.2	5.7	2.2	11.2	5.3	2.3
	$F_{SUR,SS}$	42.2	12.4	0.9	44.1	13.5	0.9	41.0	11.8	1.0
	$F_{SS,SUR}$	94.7	92.0	88.2	94.2	91.7	87.6	94.0	90.9	85.7
(3,4)	GFESM	11.9	6.4	2.5	12.0	6.2	2.5	11.9	5.4	1.8
	$F_{SUR,SS}$	42.7	12.5	1.1	43.5	13.1	1.1	41.0	11.0	1.1
	$F_{SS,SUR}$	94.3	92.3	88.5	93.4	91.4	87.0	92.9	90.8	85.9

Total number of observations $T=136$.

^aThe DGP is given in (7) with $\mu_1^* = \{-2, 2, 0, 0\}$, $\mu_2^* = \{-2, 0, 0, 2\}$, $\mu_3^* = \{0, 2, -2, 0\}$, and $\mu_4^* = \{0, 0, -2, 2\}$ for the quarters (1,2), (1,4), (2,3), and (3,4) respectively.

^bSee footnote ^a of Table 3.

Table 7
Forecast performance, when the DGP is a process with one or more seasonal unit roots^a

$\Phi(B)$	Criterion ^b	One-step ahead number of periods			Four-steps ahead number of periods			Eight-steps ahead number of periods		
		16	32	64	16	32	64	16	32	64
$(1-B)^4$	GFESM	69.3	79.8	90.0	69.9	80.0	90.2	8.6	39.4	72.8
	$F_{SUR,SS}$	94.7	93.1	91.8	88.7	87.5	87.3	76.6	74.4	73.9
	$F_{SS,SUR}$	67.8	40.5	14.4	61.8	37.4	14.2	50.8	29.8	11.1
$(1+B^2) \times (1-B)$	GFESM	63.7	72.5	83.1	64.3	72.3	87.8	7.5	32.7	62.0
	$F_{SUR,SS}$	93.4	90.8	86.8	79.6	75.4	70.0	63.6	58.5	53.3
	$F_{SS,SUR}$	73.3	50.5	24.7	63.2	42.7	22.5	49.6	34.1	17.3
$(1+B) \times (1-B)$	GFESM	70.2	76.8	84.5	70.3	76.5	84.5	6.8	28.8	58.1
	$F_{SUR,SS}$	93.5	92.7	90.1	62.2	62.1	58.0	41.8	43.9	42.4
	$F_{SS,SUR}$	76.7	57.9	32.7	49.4	40.1	25.5	31.2	26.8	16.7
$(1-B)$	GFESM	69.6	73.9	79.5	69.4	73.5	79.7	66.1	71.4	77.8
	$F_{SUR,SS}$	94.0	92.0	89.6	94.2	92.4	90.4	93.7	92.5	90.4
	$F_{SS,SUR}$	76.6	61.3	46.0	76.5	62.0	45.9	77.0	61.6	45.6

Total number of observations $T=136$.

^aThe DGP is given in (8).

^bSee footnote ^a of Table 3.

seasonal unit root instead of 2 or 3 seasonal unit roots. Notice from Table 7 that the percentage that the model with possible seasonal unit roots leads to better forecasts in case the DGP is model (8) (about 70%) is lower than the percentage that the seasonal mean shift model leads to better forecasts in case the DGP is model (7) (about 85%), see e.g. Table 3.

3.3. Several other DGPs

So far, the DGPs correspond with one of the models of interest, which are used to generate forecasts. In this subsection we consider DGPs which are different from the models under investigation. It is interesting to see how the seasonal mean shift model and the model with possible seasonal unit roots forecast if the DGP contains an alternative description of changing seasonality.

The first DGP contains several seasonal mean shifts. The location of these shifts is determined by a Markov process, i.e.,

$$\Delta_1 y_t = \sum_{s=1}^4 \mu_s D_{st} + \sum_{s=1}^4 \mu_s^* D_{st} s_t + \varepsilon_t, \quad (9)$$

$$\varepsilon_t \sim NID(0, 1),$$

where s_t follows a first order Markov process with transition probabilities $p = \Pr[s_t = 0 | s_{t-1} = 0]$ and $q = \Pr[s_t = 1 | s_{t-1} = 1]$, see Hamilton (1989). This DGP provides a simple framework to mimic a business cycle, see also Canova and Ghysels (1994).

A seasonal mean, which evolves over time as a martingale, is another possibility to generate a changing seasonal pattern, see Canova and Hansen (1995). We consider the DGP

$$\Delta_1 y_t = \mu_t (-1)^t + \varepsilon_t, \quad (10)$$

with $\mu_t = \mu_{t-1} + v_t$, $\varepsilon_t \sim NID(0, 1)$, and $v_t \sim NID(0, \sigma^2)$. Note that when $\sigma^2 = 0$, the DGP in (10) is identical to (7) with $\mu_s^* = \{0, 0, 0, 0\}$.

Unreported simulation results show that considering one- and four-step ahead forecasts, the model with possible seasonal unit roots outperforms the

model with seasonal mean shifts. For the eight-step ahead forecasts this is not the case if the probability of a regime change is small (i.e., p and q are large). If the probability of staying in the same regime decreases the forecast performance of the seasonal mean shift model deteriorates. The encompassing tests display the same pattern. The results for the DGP in (10) are roughly the same as for the DGP in (8) with $\Phi(B) = (1+B)(1-B)$. This is not surprising since the DGP in (10) resembles a process with a unit root -1 .

In summary, the Monte Carlo simulations in this section show that it is important to account for possible seasonal mean shifts if they are present. Neglecting seasonal mean shifts can have large consequences for the forecasting performance of the selected model. On the other hand, incorrectly modeling seasonal stochastic trends by a one-time seasonal mean shift also results in inferior forecasts.

One may conjecture that outcomes in the simulation exercises are biased due to the proposed model selection strategy. However, this model selection procedure reflects the way in which empirical data are often modeled. To investigate the effects of the proposed model selection strategy on our simulation results we run some experiments where we impose the correct lag order and the location of the breakpoint in our simulation experiments. These unreported simulation results show that then the percentage that the correct model produces better forecasts increases substantially.

4. Conclusion

Recent research has shown that seasonal mean shifts bias tests for seasonal unit roots towards nonrejection of seasonal unit roots. In the present paper we have addressed the choice between the seasonal mean shift model and the seasonal unit root model from a forecasting perspective. By means of simulation, we investigated the impact of imposing the incorrect model on forecasting performance. Data generating processes with seasonal mean shifts as well as with seasonal unit roots have been considered.

If a seasonal mean shift is present, the simulation

exercise shows that in about 80% to 95% of the cases the model selected by the HEGY procedure is outperformed by the seasonal mean shift model. This is especially the case if the shift does not occur near the end of the estimation sample. Results are insensitive to which of the four seasons is affected by a seasonal mean shift. However, the number of changing seasons has some impact on the results. On the other hand, if one or more seasonal unit roots are present, the model selected by the HEGY procedure outperforms the seasonal mean shift model in 70% of the cases for one- and four-step ahead forecasts. Surprisingly, the seasonal mean shift model does often significantly better for eight-step ahead forecasts.

Our simulation results show the importance of selecting the most appropriate model for forecasting purposes. This was also illustrated by considering three consumption series, which provided the motivation of our simulation exercises. Although the HEGY procedure selects a model with three seasonal unit roots for all series, the forecasts of this model are often outperformed by the model with seasonal mean shifts. In practice, when the HEGY procedure indicates the presence of one or more seasonal unit roots, it is advisable to allow for possible seasonal mean shifts in the HEGY test regressions, see Franses and Vogelsang (1997) or Franses et al. (1997). This contributes to a better knowledge about the number of seasonal unit roots, which in turn may result in a better forecast performance of the model. If the evidence for seasonal unit roots disappears, it is useful to consider models with deterministic seasonal mean shifts. However, the fact that these models are not capable of forecasting seasonal mean shifts in the out-of-sample period, can be seen as a disadvantage. If it is likely that such events can happen again in the future, it may be sensible to consider models which allow for stochastically changing seasonal means.

Acknowledgments

The research by the first author was sponsored by the Economic Research Foundation, which is part of the Netherlands Organisation for Scientific Research

(N.W.O.). The second author thanks the financial support from the Royal Netherlands Academy of Arts and Sciences. We thank three anonymous referees and Jan de Gooijer (who acted as editor) for their comments.

References

- Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.
- Canova, F., Ghysels, E., 1994. Changes in seasonal patterns: are they cyclical?. *Journal of Economic Dynamics and Control* 18, 1143–1171.
- Canova, F., Hansen, B.E., 1995. Are seasonal patterns constant over time? A test for seasonal stability. *Journal of Business & Economic Statistics* 13, 237–252.
- Clements, M.P., Hendry, D., 1993. On the limitations of comparing mean squared forecast errors. *Journal of Forecasting* 12, 617–637.
- Franses, P.H., 1996. *Periodicity and Stochastic Trends in Economic Time Series*. Oxford University Press, Oxford.
- Franses, P.H., Hoek, H., Paap, R., 1997. Bayesian analysis of seasonal unit roots and seasonal mean shifts. *Journal of Econometrics* 78, 359–380.
- Franses, P.H., Vogelsang, T.J., 1997. On seasonal cycles, unit roots, mean shifts. forthcoming in *Review of Economics and Statistics*.
- Ghysels, E., 1994. On the Economics and Econometrics of Seasonality. In: Sims, C.A. (Ed.), *Advances in Econometrics*. Cambridge University Press, Cambridge, pp. 257–316.
- Hamilton, J.D., 1989. A new approach to the econometric analysis of nonstationary time series and business cycles. *Econometrica* 57, 357–384.
- Hendry, D., 1979. Predictive Failure and Econometric Modelling in Macroeconomics. In: Ormerod, P. (Ed.), *Economic Modelling*. Heineman, London, pp. 217–242.
- Hylleberg, S., 1994. Modelling Seasonal Variation. In: Hargreaves, C.P. (Ed.), *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press, Oxford, pp. 153–178.
- Hylleberg, S., Engle, R.F., Granger, C.W.J., Yoo, B., 1990. Seasonal integration and cointegration. *Journal of Econometrics* 44, 215–238.
- Hylleberg, S., Jørgensen, C., Sørensen, N.K., 1993. Seasonality in macroeconomic time series. *Empirical Economics* 18, 321–335.
- Osborn, D.R., 1988. Seasonality and habit persistence in a life cycle model of consumption. *Journal of Applied Econometrics* 3, 255–266.
- Osborn, D.R., 1990. A survey of seasonality in UK macroeconomic variables. *International Journal of Forecasting* 6, 327–336.

Otto, G., Wirjanto, T., 1990. Seasonal unit roots tests on Canadian macroeconomic time series. *Economic Letters* 34, 117–120.

Perron, P., 1990. Testing for a unit root in a time series with a changing mean. *Journal of Business & Economic Statistics* 8, 153–162.

Biographies: Richard PAAP is PhD student at the Tinbergen Institute, Erasmus University Rotterdam. His research topics are Bayesian and classical time series analysis, cointegration, seasonality and nonlinearity.

Philip Hans FRANSES is Associate Professor in Econometrics and the Director of the Rotterdam Institute for Business Economic Studies, both at the Erasmus University Rotterdam. His research interests include time series forecasting, seasonality and non-linearity.

Henk HOEK is PhD student at the Tinbergen Institute, Erasmus University Rotterdam. His research topic is Bayesian analysis of economic time series.