Chapter 13

FORECASTING SEASONAL TIME SERIES

ERIC GHYSELS

Department of Economics, University of North Carolina

DENISE R. OSBORN

School of Economic Studies, University of Manchester

PAULO M.M. RODRIGUES

Faculty of Economics, University of Algarve

Contents

Handbook of Economic Forecasting, Volume 1
Edited by Graham Elliott, Clive W.J. Granger and Allan Timmermann

DOI: 10.1016/S1574-0706(05)01013-X

Abstract

This chapter reviews the principal methods used by researchers when forecasting seasonal time series. In addition, the often overlooked implications of forecasting and feedback for seasonal adjustment are discussed. After an introduction in Section 1, Section 2 examines traditional univariate linear models, including methods based on SARIMA models, seasonally integrated models and deterministic seasonality models. As well as examining how forecasts are computed in each case, the forecast implications of misspecifying the class of model (deterministic versus nonstationary stochastic) are considered. The linear analysis concludes with a discussion of the nature and implications of cointegration in the context of forecasting seasonal time series, including merging short-term seasonal forecasts with those from long-term (nonseasonal) models.

Periodic (or seasonally varying parameter) models, which often arise from theoretical models of economic decision-making, are examined in Section 3. As periodic models may be highly parameterized, their value for forecasting can be open to question. In this context, modelling procedures for periodic models are critically examined, as well as procedures for forecasting.

Section 3 discusses less traditional models, specifically nonlinear seasonal models and models for seasonality in variance. Such nonlinear models primarily concentrate on interactions between seasonality and the business cycle, either using a threshold specification to capture changing seasonality over the business cycle or through regime transition probabilities being seasonally varying in a Markov switching framework. Seasonality heteroskedasticity is considered for financial time series, including deterministic versus stochastic seasonality, periodic GARCH and periodic stochastic volatility models for daily or intra-daily series.

Economists typically consider that seasonal adjustment rids their analysis of the "nuisance" of seasonality. Section 5 shows this to be false. Forecasting seasonal time series is an inherent part of seasonal adjustment and, further, decisions based on seasonally adjusted data affect future outcomes, which destroys the assumed orthogonality between seasonal and nonseasonal components of time series.

Keywords

seasonality, seasonal adjustment, forecasting with seasonal models, nonstationarity, nonlinearity, seasonal cointegration models, periodic models, seasonality in variance

JEL classification: C22, C32, C53

1. Introduction

Although seasonality is a dominant feature of month-to-month or quarter-to-quarter fluctuations in economic time series [Beaulieu and Miron (1992), Miron (1996), Franses (1996)], it has typically been viewed as of limited interest by economists, who generally use seasonally adjusted data for modelling and forecasting. This contrasts with the perspective of the economic agent, who makes (say) production or consumption decisions in a seasonal context [Ghysels (1988, 1994a), Osborn (1988)].

In this chapter, we study forecasting of seasonal time series and its impact on seasonal adjustment. The bulk of our discussion relates to the former issue, where we assume that the (unadjusted) value of a seasonal series is to be forecast, so that modelling the seasonal pattern itself is a central issue. In this discussion, we view seasonal movements as an inherent feature of economic time series which should be integrated into the econometric modelling and forecasting exercise. Hence, we do not consider seasonality as a separable component in the unobserved components methodology, which is discussed in Chapter 7 in this Handbook [see Harvey (2006)]. Nevertheless, such unobserved components models do enter our discussion, since they are the basis of official seasonal adjustment. Our focus is then not on the seasonal models themselves, but rather on how forecasts of seasonal time series enter the adjustment process and, consequently, influence subsequent decisions. Indeed, the discussion here reinforces our position that seasonal and nonseasonal components are effectively inseparable.

Seasonality is the periodic and largely repetitive pattern that is observed in time series data over the course of a year. As such, it is largely predictable. A generally agreed definition of seasonality in the context of economics is provided by Hylleberg (1992, p. 4) as follows: "*Seasonality is the systematic, although not necessarily regular, intra-year movement caused by the changes of weather, the calendar, and timing of decisions, directly or indirectly through the production and consumption decisions made by the agents of the economy. These decisions are influenced by endowments, the expectations and preferences of the agents, and the production techniques available in the economy*." This definition implies that seasonality is not necessarily fixed over time, despite the fact that the calendar does not change. Thus, for example, the impact of Christmas on consumption or of the summer holiday period on production may evolve over time, despite the timing of Christmas and the summer remaining fixed.

Intra-year observations on most economic time series are typically available at quarterly or monthly frequencies, so our discussion concentrates on these frequencies. We follow the literature in referring to each intra-year observation as relating to a "season", by which we mean an individual month or quarter. Financial time series are often observed at higher frequencies, such as daily or hourly and methods analogous to those discussed here can be applied when forecasting the patterns of financial time series that are associated with the calendar, such as days of the week or intradaily patterns. However, specific issues arise in forecasting financial time series, which is not the topic of the present chapter.

In common with much of the forecasting literature, our discussion assumes that the forecaster aims to minimize the mean-square forecast error (MSFE). As shown by Whittle (1963) in a linear model context, the optimal (minimum MSFE) forecast is given by the expected value of the future observation y_{T+h} conditional on the information set, $y_1, \ldots, y_T$, available at time T, namely

$$\widehat{y}_{T+h|T} = E(y_{T+h}|y_1, \ldots, y_T). \tag{1}$$

However, the specific form of $\widehat{y}_{T+h|T}$ depends on the model assumed to be the data generating process (DGP).

When considering the optimal forecast, the treatment of seasonality may be expected to be especially important for short-run forecasts, more specifically forecasts for horizons h that are less than one year. Denoting the number of observations per year as S, then this points to $h = 1, \ldots, S - 1$ as being of particular interest. Since $h = S$ is a one-year ahead forecast, and seasonality is typically irrelevant over the horizon of a year, seasonality may have a smaller role to play here than at shorter horizons. Seasonality obviously once again comes into play for horizons $h = S + 1, \ldots, 2S - 1$ and at subsequent horizons that do not correspond to an integral number of years.

Nevertheless, the role of seasonality should not automatically be ignored for forecasts at horizons of an integral number of years. If seasonality is changing, then a model that captures this changing seasonal pattern should yield more accurate forecasts at these horizons than one that ignores it.

This chapter is structured as follows. In Section 2 we briefly introduce the widely-used classes of univariate SARIMA and deterministic seasonality models and show how these are used for forecasting purposes. Moreover, an analysis on forecasting with misspecified seasonal models is presented. This section also discusses Seasonal Cointegration, including the use of Seasonal Cointegration Models for forecasting purposes, and presents the main conclusions of forecasting comparisons that have appeared in the literature. The idea of merging short- and long-run forecasts, put forward by Engle, Granger and Hallman (1989), is also discussed.

Section 3 discusses the less familiar periodic models where parameters change over the season; such models often arise from economic theories in a seasonal context. We analyze forecasting with these models, including the impact of neglecting periodic parameter variation and we discuss proposals for more parsimonious periodic specifications that may improve forecast accuracy. Periodic cointegration is also considered and an overview of the few existing results of forecast performance of periodic models is presented.

In Section 4 we move to recent developments in modelling seasonal data, specifically nonlinear seasonal models and models that account for seasonality in volatility. Nonlinear models include those of the threshold and Markov switching types, where the focus is on capturing business cycle features in addition to seasonality in the conditional mean. On the other hand, seasonality in variance is important in finance; for instance, Martens, Chang and Taylor (2002) show that explicitly modelling intraday seasonality improves out-of-sample forecasting performance.

The final substantive section of this chapter turns to the interactions of seasonality and seasonal adjustment, which is important due to the great demand for seasonally adjusted data. This section demonstrates that such adjustment is not separable from forecasting the seasonal series. Further, we discuss the feedback from seasonal adjustment to seasonality that exists when the actions of policymakers are considered.

In addition to general conclusions, Section 6 draws some implications from the chapter that are relevant to the selection of a forecasting model in a seasonal context.

2. Linear models

Most empirical models applied when forecasting economic time series are linear in parameters, for which the model can be written as

$$y_{Sn+s} = \mu_{Sn+s} + x_{Sn+s}, \tag{2}$$

$$\phi(L)x_{Sn+s} = u_{Sn+s} \tag{3}$$

where y_{Sn+s} ($s = 1, \ldots, S, n = 0, \ldots, T-1$) represents the observable variable in season (e.g., month or quarter) s of year n, the polynomial $\phi(L)$ contains any unit roots in y_{Sn+s} and will be specified in the following subsections according to the model being discussed, L represents the conventional lag operator, $L^k x_{Sn+s} \equiv x_{Sn+s-k}$, $k = 0, 1, \ldots$, the driving shocks $\{u_{Sn+s}\}$ of (3) are assumed to follow an ARMA(p, q), $0 \leqslant p, q < \infty$ process, such as, $\beta(L)u_{Sn+s} = \theta(L)\varepsilon_{Sn+s}$, where the roots of $\beta(z) \equiv 1 - \sum_{j=1}^{p} \beta_j z^j = 0$ and $\theta(z) \equiv 1 - \sum_{j=1}^{q} \theta_j z^j = 0$ lie outside the unit circle, $|z| = 1$, with $\varepsilon_{Sn+s} \sim \text{iid}(0, \sigma^2)$. The term μ_{Sn+s} represents a deterministic kernel which will be assumed to be either (i) a set of seasonal means, i.e., $\sum_{s=1}^{S} \delta_s D_{s,Sn+s}$ where $D_{i,Sn+s}$ is a dummy variable taking value 1 in season i and zero elsewhere, or (ii) a set of seasonals with a (nonseasonal) time trend, i.e., $\sum_{s=1}^{S} \delta_s D_{s,Sn+s} + \tau(Sn+s)$. In general, the second of these is more plausible for economic time series, since it allows the underlying level of the series to trend over time, whereas $\mu_{Sn+s} = \delta_s$ implies a constant underlying level, except for seasonal variation.

When considering forecasts, we use T to denote the total (observed) sample size, with forecasts required for the future period $T + h$ for $h = 1, 2, \ldots$.

Linear seasonal forecasting models differ essentially in their assumptions about the presence of unit roots in $\phi(L)$. The two most common forms of seasonal models in empirical economics are seasonally integrated models and models with deterministic seasonality. However, seasonal autoregressive integrated moving average (SARIMA) models retain an important role as a forecasting benchmark. Each of these three models and their associated forecasts are discussed in a separate subsection below.

2.1. SARIMA model

When working with nonstationary seasonal data, both annual changes and the changes between adjacent seasons are important concepts. This motivated Box and Jenkins

(1970) to propose the SARIMA model

$$\beta(L)(1-L)\big(1-L^S\big)y_{Sn+s} = \theta(L)\varepsilon_{Sn+s} \tag{4}$$

which results from specifying $\phi(L) = \Delta_1\Delta_S = (1-L)(1-L^S)$ in (3). It is worth noting that the imposition of $\Delta_1\Delta_S$ annihilates the deterministic variables (seasonal means and time trend) of (2), so that these do not appear in (4). The filter $(1-L^S)$ captures the tendency for the value of the series for a particular season to be highly correlated with the value for the same season a year earlier, while $(1-L)$ can be motivated as capturing the nonstationary nonseasonal stochastic component. This model is often found in textbooks, see for instance Brockwell and Davis (1991, pp. 320–326) and Harvey (1993, pp. 134–137). Franses (1996, pp. 42–46) fits SARIMA models to various real macroeconomic time series.

An important characteristic of model (4) is the imposition of unit roots at all seasonal frequencies, as well as two unit roots at the zero frequency. This occurs as $(1-L)(1-L^S) = (1-L)^2(1+L+L^2+\cdots+L^{S-1})$, where $(1-L)^2$ relates to the zero frequency while the moving annual sum $(1+L+L^2+\cdots+L^{S-1})$ implies unit roots at the seasonal frequencies (see the discussion below for seasonally integrated models). However, the empirical literature does not provide much evidence favoring the presence of two zero frequency unit roots in observed time series [see, e.g., Osborn (1990) and Hylleberg, Jørgensen and Sørensen (1993)], which suggests that the SARIMA model is overdifferenced. Although these models may seem empirically implausible, they can be successful in forecasting due to their parsimonious nature.

More specifically, the special case of (4) where

$$(1-L)\big(1-L^S\big)y_{Sn+s} = (1-\theta_1 L)\big(1-\theta_S L^S\big)\varepsilon_{Sn+s} \tag{5}$$

with $|\theta_1| < 1$, $|\theta_S| < 1$ retains an important position. This is known as the airline model because Box and Jenkins (1970) found it appropriate for monthly airline passenger data. Subsequently, the model has been shown to provide robust forecasts for many observed seasonal time series, and hence it often provides a benchmark for forecast accuracy comparisons.

2.1.1. Forecasting with SARIMA models

Given that ε_{T+h} is assumed to be iid$(0, \sigma^2)$, and if all parameters are known, the optimal (minimum MSFE) h-step ahead forecast of $\Delta_1\Delta_S y_{T+h}$ for the airline model (5) is, from (1),

$$\begin{aligned}\Delta_1\Delta_S\widehat{y}_{T+h|T} = {}& -\theta_1 E(\varepsilon_{T+h-1}|y_1,\ldots,y_T) - \theta_S E(\varepsilon_{T+h-S}|y_1,\ldots,y_T)\\ & + \theta_1\theta_S E(\varepsilon_{T+h-S-1}|y_1,\ldots,y_T), \quad h \geqslant 1\end{aligned} \tag{6}$$

where $E(\varepsilon_{T+h-i}|y_1,\ldots,y_T) = 0$ if $h > i$ and $E(\varepsilon_{T+h-i}|y_1,\ldots,y_T) = \varepsilon_{T+h-i}$ if $h \leqslant i$. Corresponding expressions can be derived for forecasts from other ARIMA models. In practice, of course, estimated parameters are used in generating these forecast values.

Forecasts of y_{T+h} for a SARIMA model can be obtained from the identity

$$\begin{aligned}E(y_{T+h}|y_1,\dots,y_T) &= E(y_{T+h-1}|y_1,\dots,y_T) + E(y_{T+h-S}|y_1,\dots,y_T) \\ &\quad - E(y_{T+h-S-1}|y_1,\dots,y_T) + \Delta_1\Delta_S\widehat{y}_{T+h|T}. \end{aligned} \tag{7}$$

Clearly, $E(y_{T+h-i}|y_1,\dots,y_T) = y_{T+h-i}$ for $h \leqslant i$, and forecasts $E(y_{T+h-i}|y_1,\dots,y_T)$ for $h > i$ required on the right-hand side of (7) can be generated recursively for $h = 1, 2, \dots$.

In this linear model context, optimal forecasts of other linear transformations of y_{T+h} can be obtained from these; for example, $\Delta_1\widehat{y}_{T+h} = \widehat{y}_{T+h} - \widehat{y}_{T+h-1}$ and $\Delta_S\widehat{y}_{T+h} = \widehat{y}_{T+h} - \widehat{y}_{T+h-S}$. In the special case of the airline model, (6) implies that $\Delta_1\Delta_S\widehat{y}_{T+h|T} = 0$ for $h > S+1$, and hence $\Delta_1\widehat{y}_{T+h|T} = \Delta_1\widehat{y}_{T+h-S|T}$ and $\Delta_S\widehat{y}_{T+h|T} = \Delta_S\widehat{y}_{T+h-1|T}$ at these horizons; see also Clements and Hendry (1997) and Osborn (2002). Therefore, when applied to forecasts for $h > S+1$, the airline model delivers a "same change" forecast, both when considered over a year and also over a single period compared to the corresponding period of the previous year.

2.2. Seasonally integrated model

Stochastic seasonality can arise through the stationary ARMA components $\beta(L)$ and $\theta(L)$ of u_{Sn+s} in (3). The case of stationary seasonality is treated in the next subsection, in conjunction with deterministic seasonality. Here we examine nonstationary stochastic seasonality where $\phi(L) = 1 - L^S = \Delta_S$ in (2). However, in contrast to the SARIMA model, the seasonally integrated model imposes only a single unit root at the zero frequency. Application of annual differencing to (2) yields

$$\beta(L)\Delta_S y_{Sn+s} = \beta(1)S\tau + \theta(L)\varepsilon_{Sn+s} \tag{8}$$

since $\Delta_S\mu_{Sn+s} = S\tau$. Thus, the seasonally integrated process of (8) has a common annual drift, $\beta(1)S\tau$, across seasons. Notice that the underlying seasonal means μ_{Sn+s} are not observed, since the seasonally varying component $\sum_{s=1}^{S}\delta_s D_{s,Sn+s}$ is annihilated by seasonal (that is, annual) differencing. In practical applications in economics, it is typically assumed that the stochastic process is of the autoregressive form, so that $\theta(L) = 1$.

As a result of the influential work of Box and Jenkins (1970), seasonal differencing has been a popular approach when modelling and forecasting seasonal time series. Note, however, that a time series on which seasonal differencing $(1 - L^S)$ needs to be applied to obtain stationarity has S roots on the unit circle. This can be seen by factorizing $(1 - L^S)$ into its evenly spaced roots, $\mathrm{e}^{\pm\mathrm{i}(2\pi k/S)}$ $(k = 0, 1, \dots, S-1)$ on the unit circle, that is, $(1-L^S) = (1-L)(1+L)\prod_{k=1}^{S^*}(1-2\cos\eta_k L+L^2) = (1-L)(1+L+\dots+L^{S-1})$ where $S^* = \mathrm{int}[(S-1)/2]$, int[.] is the integer part of the expression in brackets and $\eta_k \in (0, \pi)$. The real positive unit root, $+1$, relates to the long-run or zero frequency, and hence is often referred to as nonseasonal, while the remaining $(S-1)$ roots represent seasonal unit roots that occur at frequencies η_k (the unit root at frequency π is known as the Nyquist frequency root and the complex roots as the harmonics). A seasonally integrated process y_{Sn+s} has unbounded spectral density at each seasonal frequency due to the presence of these unit roots.

From an economic point of view, nonstationary seasonality can be controversial because the values over different seasons are not cointegrated and hence can move in any direction in relation to each other, so that "*winter can become summer*". This appears to have been first noted by Osborn (1993). Thus, the use of seasonal differences, as in (8) or through the multiplicative filter as in (4), makes rather strong assumptions about the stochastic properties of the time series under analysis. It has, therefore, become common practice to examine the nature of the stochastic seasonal properties of the data via seasonal unit root tests. In particular, Hylleberg, Engle, Granger and Yoo [HEGY] (1990) propose a test for the null hypothesis of seasonal integration in quarterly data, which is a seasonal generalization of the Dickey–Fuller [DF] (1979) test. The HEGY procedure has since been extended to the monthly case by Beaulieu and Miron (1993) and Taylor (1998), and was generalized to any periodicity S, by Smith and Taylor (1999).[1]

2.2.1. Testing for seasonal unit roots

Following HEGY and Smith and Taylor (1999), *inter alia*, the regression-based approach to testing for seasonal unit roots implied by $\phi(L) = 1 - L^S$ can be considered in two stages. First, the OLS de-meaned series $\widetilde{x}_{Sn+s} = y_{Sn+s} - \hat{\mu}_{Sn+s}$ is obtained, where $\hat{\mu}_{Sn+s}$ is the fitted value from the OLS regression of y_{Sn+s} on an appropriate set of deterministic variables. Provided μ_{Sn+s} is not estimated under an overly restrictive case, the resulting unit root tests will be exact invariant to the parameters characterizing the mean function μ_{Sn+s}; see Burridge and Taylor (2001).

Following Smith and Taylor (1999), $\phi(L)$ in (3) is then linearized around the seasonal unit roots $\exp(\pm \mathrm{i}2\pi k/S)$, $k = 0, \ldots, [S/2]$, so that the auxiliary regression equation

$$\begin{aligned}\Delta_S \widetilde{x}_{Sn+s} &= \pi_0 \widetilde{x}_{0,Sn+s-1} + \pi_{S/2} \widetilde{x}_{S/2,Sn+s-1} \\ &\quad + \sum_{k=1}^{S^*} \left(\pi_{\alpha,k} \widetilde{x}^{\alpha}_{k,Sn+s-1} + \pi_{\beta,k} \widetilde{x}^{\beta}_{k,Sn+s-1}\right) \\ &\quad + \sum_{j=1}^{p^*} \beta_j^* \Delta_S \widetilde{x}_{Sn+s-j} + \varepsilon_{Sn+s} \end{aligned} \tag{9}$$

is obtained. The regressors are linear transformations of $\widetilde{x}_{Sn+s}$, namely

$$\widetilde{x}_{0,Sn+s} \equiv \sum_{j=0}^{S-1} \widetilde{x}_{Sn+s-j}, \qquad \widetilde{x}_{S/2,Sn+s} \equiv \sum_{j=0}^{S-1} \cos\big[(j+1)\pi\big] \widetilde{x}_{Sn+s-j},$$

[1] Numerous other seasonal unit root tests have been developed; see *inter alia* Breitung and Franses (1998), Busetti and Harvey (2003), Canova and Hansen (1995), Dickey, Hasza and Fuller (1984), Ghysels, Lee and Noh (1994), Hylleberg (1995), Osborn et al. (1988), Rodrigues (2002), Rodrigues and Taylor (2004a, 2004b) and Taylor (2002, 2003). However, in practical applications, the HEGY test is still the most widely applied.

$$\widetilde{x}^{\alpha}_{k,Sn+s} \equiv \sum_{j=0}^{S-1} \cos\bigl[(j+1)\omega_k\bigr]\widetilde{x}_{Sn+s-j},$$
$$\widetilde{x}^{\beta}_{k,Sn+s} \equiv -\sum_{j=0}^{S-1} \sin\bigl[(j+1)\omega_k\bigr]\widetilde{x}_{Sn+s-j}, \tag{10}$$

with $k = 1, \ldots, S^*$, $S^* = \operatorname{int}[(S-1)/2]$. For example, in the quarterly case, $S = 4$, the relevant transformations are:

$$\widetilde{x}_{0,Sn+s} \equiv \bigl(1 + L + L^2 + L^3\bigr)\widetilde{x}_{Sn+s}, \qquad \widetilde{x}_{2,Sn+s} \equiv -\bigl(1 - L + L^2 - L^3\bigr)\widetilde{x}_{Sn+s},$$
$$\widetilde{x}^{\alpha}_{1,Sn+s} \equiv \widetilde{x}_{1,Sn+s-1} = -L\bigl(1 - L^2\bigr)\widetilde{x}_{Sn+s},$$
$$\widetilde{x}^{\beta}_{1,Sn+s} \equiv \widetilde{x}_{1,Sn+s} = -\bigl(1 - L^2\bigr)\widetilde{x}_{Sn+s}. \tag{11}$$

The regression (9) can be estimated over observations $Sn + s = p^* + S + 1, \ldots, T$, with $\pi_{S/2}\widetilde{x}_{S/2,Sn+s-1}$ omitted if S is odd. Note also that the autoregressive order p^* used must be sufficiently large to satisfactorily account for any autocorrelation, including any moving average component in (8).

The presence of unit roots implies exclusion restrictions for π_0, $\pi_{k,\alpha}$, $\pi_{k,\beta}$, $k = 1, \ldots, S^*$, and $\pi_{S/2}$ (S even), while the overall null hypothesis of seasonal integration implies all these are zero. To test seasonal integration against stationarity at one or more of the seasonal or nonseasonal frequencies, HEGY suggest using: t_0 (left-sided) for the exclusion of $\widetilde{x}_{0,Sn+s-1}$; $t_{S/2}$ (left-sided) for the exclusion of $\widetilde{x}_{S/2,Sn+s-1}$ (S even); F_k for the exclusion of *both* $\widetilde{x}^{\alpha}_{k,Sn+s-1}$ and $\widetilde{x}^{\beta}_{k,Sn+s-1}$, $k = 1, \ldots, S^*$. These tests examine the potential unit roots separately at each of the zero and seasonal frequencies, raising issues of the significance level for the overall test (Dickey, 1993). Consequently, Ghysels, Lee and Noh (1994), also consider joint frequency OLS F-statistics. Specifically $F_{1\ldots[S/2]}$ tests for the presence of all seasonal unit roots by testing for the exclusion of $\widetilde{x}_{S/2,Sn+s-1}$ (S even) and $\{\widetilde{x}^{\alpha}_{k,Sn+s-1}, \widetilde{x}^{\beta}_{k,Sn+s-1}\}_{k=1}^{S^*}$, while $F_{0\ldots[S/2]}$ examines the overall null hypothesis of seasonal integration, by testing for the exclusion of $\widetilde{x}_{0,Sn+s-1}$, $\widetilde{x}_{S/2,Sn+s-1}$ (S even), and $\{\widetilde{x}^{\alpha}_{k,Sn+s-1}, \widetilde{x}^{\beta}_{k,Sn+s-1}\}_{k=1}^{S^*}$ in (9). These joint tests are further considered by Taylor (1998) and Smith and Taylor (1998, 1999).

Empirical evidence regarding seasonal integration in quarterly data is obtained by (among others) HEGY, Lee and Siklos (1997), Hylleberg, Jørgensen and Sørensen (1993), Mills and Mills (1992), Osborn (1990) and Otto and Wirjanto (1990). The monthly case has been examined relatively infrequently, but relevant studies include Beaulieu and Miron (1993), Franses (1991) and Rodrigues and Osborn (1999). Overall, however, there is little evidence that the seasonal properties of the data justify application of the Δ_s filter for economic time series. Despite this, Clements and Hendry (1997) argue that the seasonally integrated model is useful for forecasting, because the seasonal differencing filter makes the forecasts robust to structural breaks in seasonality.[2]

[2] Along slightly different lines it is also worth noting that Ghysels and Perron (1996) show that traditional seasonal adjustment filters also mask structural breaks in nonseasonal patterns.

On the other hand, Kawasaki and Franses (2004) find that imposing individual seasonal unit roots on the basis of model selection criteria generally improves one-step ahead forecasts for monthly industrial production in OECD countries.

2.2.2. Forecasting with seasonally integrated models

As they are linear, forecasts from seasonally integrated models are generated in an analogous way to SARIMA models. Assuming all parameters are known and there is no moving average component (i.e., $\theta(L) = 1$), the optimal forecast is given by

$$\begin{aligned}\Delta_S \widehat{y}_{T+h|T} &= \beta(1) S\tau + \sum_{i=1}^{p} \beta_i E(\Delta_S y_{T+h-i} | y_1, \ldots, y_T) \\ &= \beta(1) S\tau + \sum_{i=1}^{p} \beta_i \Delta_S \widehat{y}_{T+h-i|T} \end{aligned} \tag{12}$$

where $\Delta_S \widehat{y}_{T+h-i|T} = \widehat{y}_{T+h-i|T} - \widehat{y}_{T+h-i-S|T}$ and $\widehat{y}_{T+h-S|T} = y_{T+h-S}$ for $h - S \leqslant 0$, with forecasts generated recursively for $h = 1, 2, \ldots$.

As noted by Ghysels and Osborn (2001) and Osborn (2002, p. 414), forecasts for other transformations can be easily obtained. For instance, the level and first difference forecasts can be derived as

$$\widehat{y}_{T+h|T} = \Delta_S \widehat{y}_{T+h|T} + \widehat{y}_{T-S+h|T} \tag{13}$$

and

$$\begin{aligned}\Delta_1 \widehat{y}_{T+h|T} &= \widehat{y}_{T+h|T} - \widehat{y}_{T+h-1|T} \\ &= \Delta_S \widehat{y}_{T+h} - (\Delta_1 \widehat{y}_{T+h-1} + \cdots + \Delta_1 \widehat{y}_{T+h-(S-1)}),\end{aligned} \tag{14}$$

respectively.

2.3. Deterministic seasonality model

Seasonality has often been perceived as a phenomenon that generates peaks and troughs within a particular season, year after year. This type of effect is well described by deterministic variables leading to what is conventionally referred to as *deterministic seasonality*. Thus, models frequently encountered in applied economics often explicitly allow for seasonal means. Assuming the stochastic component x_{Sn+s} of y_{Sn+s} is stationary, then $\phi(L) = 1$ and (2)/(3) implies

$$\beta(L) y_{Sn+s} = \sum_{i=1}^{S} \beta(L) \mu_{Sn+s} + \theta(L) \varepsilon_{Sn+s} \tag{15}$$

where ε_{Sn+s} is again a zero mean white noise process. For simplicity of exposition, and in line with usual empirical practice, we assume the absence of moving average components, i.e., $\theta(L) = 1$. Note, however, that stationary stochastic seasonality may also enter through $\beta(L)$.

Although the model in (15) assumes a stationary stochastic process, it is common, for most economic time series, to find evidence favouring a zero frequency unit root. Then $\phi(L) = 1 - L$ plays a role and the deterministic seasonality model is

$$\beta(L)\Delta_1 y_{Sn+s} = \sum_{s=1}^{S} \beta(L)\Delta_1 \mu_{Sn+s} + \varepsilon_{Sn+s} \tag{16}$$

where $\Delta_1 \mu_{Sn+s} = \mu_{Sn+s} - \mu_{Sn+s-1}$, so that (only) the change in the seasonal mean is identified.

Seasonal dummies are frequently employed in empirical work within a linear regression framework to represent seasonal effects [see, for example, Barsky and Miron (1989), Beaulieu, Mackie-Mason and Miron (1992), and Miron (1996)]. One advantage of considering seasonality as deterministic lies in the simplicity with which it can be handled. However, consideration should be given to various potential problems that can occur when treating a seasonal pattern as purely deterministic. Indeed, spurious deterministic seasonality emerges when seasonal unit roots present in the data are neglected [Abeysinghe (1991, 1994), Franses, Hylleberg and Lee (1995), and Lopes (1999)]. On the other hand, however, Ghysels, Lee and Siklos (1993) and Rodrigues (2000) establish that, for some purposes, (15) or (16) can represent a valid approach even with seasonally integrated data, provided the model is adequately augmented to take account of any seasonal unit roots potentially present in the data.

The core of the deterministic seasonality model is the seasonal mean effects, namely μ_{Sn+s} and $\Delta_1 \mu_{Sn+s}$, for (15) and (16), respectively. However, there are a number of (equivalent) different ways that these may be represented, whose usefulness depends on the context. Therefore, we discuss this first. For simplicity, we assume the form of (15) is used and refer to μ_{Sn+s}. However, corresponding comments apply to $\Delta_1 \mu_{Sn+s}$ in (16).

2.3.1. Representations of the seasonal mean

When $\mu_{Sn+s} = \sum_{s=1}^{S} \delta_s D_{s,Sn+s}$, the mean relating to each season is constant over time, with $\mu_{Sn+s} = \mu_s = \delta_s$ ($n = 1, 2, \ldots, s = 1, 2, \ldots, S$). This is a conditional mean, in the sense that $\mu_{Sn+s} = E[y_{Sn+s} | t = Sn + s]$ depends on the season s. Since all seasons appear with the same frequency over a year, the corresponding unconditional mean is $E(y_{Sn+s}) = \mu = (1/S)\sum_{s=1}^{S} \mu_s$. Although binary seasonal dummy variables, $D_{s,Sn+s}$, are often used to capture the seasonal means, this form has the disadvantage of not separately identifying the unconditional mean of the series.

Equivalently to the conventional representation based on $D_{s,Sn+s}$, we can identify the unconditional mean through the representation

$$\mu_{Sn+s} = \mu + \sum_{s=1}^{S} \delta_s^* D_{s,Sn+s}^* \tag{17}$$

where the dummy variables $D^*_{s,Sn+s}$ are constrained to sum to zero over the year, $\sum_{s=1}^{S} D^*_{s,Sn+s} = 0$. To avoid exact multicollinearity, only $S - 1$ such dummy variables can be included, together with the intercept, in a regression context. The constraint that these variables sum to zero then implies the parameter restriction $\sum_{s=1}^{S} \delta^*_s = 0$, from which the coefficient on the omitted dummy variable can be retrieved. One specific form of such dummies is the so-called centered seasonal dummy variables, which are defined as $D^*_{s,Sn+s} = D_{s,Sn+s} - (1/S)\sum_{s=1}^{S} D_{s,Sn+s}$.[3] Nevertheless, care in interpretation is necessary in (17), as the interpretation of δ^*_s depends on the definition of $D^*_{s,Sn+s}$. For example, the coefficients of $D^*_{s,Sn+s} = D_{s,Sn+s} - (1/S)\sum_{s=1}^{S} D_{s,Sn+s}$ do not have a straightforward seasonal mean deviation interpretation.

A specific form sometimes used for (17) relates the dummy variables to the seasonal frequencies considered above for seasonally integrated models, resulting in the trigonometric representation [see, for example, Harvey (1993, 1994), or Ghysels and Osborn (2001)]

$$\mu_{Sn+s} = \mu + \sum_{j=1}^{S^{**}} \left(\gamma_j \cos \lambda_{jSn+s} + \gamma^*_j \sin \lambda_{jSn+s}\right) \tag{18}$$

where $S^{**} = \text{int}[S/2]$, and $\lambda_{jt} = \frac{2\pi j}{S}$, $j = 1, \ldots, S^{**}$. When S is even, the sine term is dropped for $j = S/2$; the number of trigonometric coefficients (γ_j, γ^*_j) is always $S - 1$.

The above comments carry over to the case when a time trend is included. For example, the use of dummies which are restricted to sum to zero with a (constant) trend implies that we can write

$$\mu_{Sn+s} = \mu + \tau(Sn + s) + \sum_{s=1}^{S} \delta^*_s D^*_{s,Sn+s} \tag{19}$$

with unconditional overall mean $E(y_{Sn+s}) = \mu + \tau(Sn + s)$.

2.3.2. Forecasting with deterministic seasonal models

Due to the prevalence of nonseasonal unit roots in economic time series, consider the model of (16), which has forecast function for $\widehat{y}_{T+h|T}$ given by

$$\widehat{y}_{T+h|T} = \widehat{y}_{T+h-1|T} + \beta(1)\tau + \sum_{i=1}^{S} \beta(L)\Delta_1 \delta_i D_{iT+h} + \sum_{j=1}^{p} \beta_j \Delta_1 \widehat{y}_{T+h-j|T} \tag{20}$$

when $\mu_{Sn+s} = \sum_{s=1}^{S} \delta_s D_{s,Sn+s} + \tau(Sn + s)$, and, as above, $\widehat{y}_{T+h-i|T} = y_{T+h-i|T}$ for $h < i$. Once again, forecasts are calculated recursively for $h = 1, 2, \ldots$ and since the

[3] These centered seasonal dummy variables are often offered as an alternative representation to conventional zero/one dummies in time series computer packages, including RATS and PcFiml.

model is linear, forecasts of other linear functions, such as $\Delta_S \widehat{y}_{T+h|T}$ can be obtained using forecast values from (20).

With $\beta(L) = 1$ and assuming $T = NS$ for simplicity, the forecast function for y_{T+h} obtained from (20) is

$$\widehat{y}_{T+h|T} = y_T + h\tau + \sum_{i=1}^{h} (\delta_i - \delta_{i-1}). \tag{21}$$

When h is a multiple of S, it is easy to see that deterministic seasonality becomes irrelevant in this expression, because the change in a purely deterministic seasonal pattern over a year is necessarily zero.

2.4. Forecasting with misspecified seasonal models

From the above discussion, it is clear that various linear models have been proposed, and are widely used, to forecast seasonal time series. In this subsection we consider the implications of using each of the three forecasting models presented above when the true DGP is a seasonal random walk or a deterministic seasonal model. These DGPs are considered because they are the simplest processes which encapsulate the key notions of nonstationary stochastic seasonality and deterministic seasonality. We first present some analytical results for forecasting with misspecified models, followed by the results of a Monte Carlo analysis.

2.4.1. Seasonal random walk

The seasonal random walk DGP is

$$y_{Sn+s} = y_{S(n-1)+s} + \varepsilon_{Sn+s}, \quad \varepsilon_{Sn+s} \sim \mathrm{iid}\big(0, \sigma^2\big). \tag{22}$$

When this seasonally integrated model is correctly specified, the one-step ahead MSFE is $E[(y_{T+1} - \widehat{y}_{T+1|T})^2] = E[(y_{T+1-S} + \varepsilon_{T+1} - y_{T+1-S})^2] = \sigma^2$.

Consider, however, applying the deterministic seasonality model (16), where the zero frequency nonstationarity is recognized and modelling is undertaken after first differencing. The relevant DGP (22) has no trend, and hence we specify $\tau = 0$. Assume a researcher naively applies the model $\Delta_1 y_{Sn+s} = \sum_{i=1}^{S} \Delta_1 \delta_i D_{i,Sn+s} + \upsilon_{Sn+s}$ with no augmentation, but (wrongly) assumes υ to be iid. Due to the presence of nonstationary stochastic seasonality, the estimated dummy variable coefficients do not asymptotically converge to constants. Although analytical results do not appear to have been derived for the resulting forecasts, we anticipate that the MSFE will converge to a degenerate distribution due to neglected nonstationarity.

On the other hand, if the dynamics are adequately augmented, then serial correlation is accounted for and the consistency of the parameter estimates is guaranteed. More specifically, the DGP (22) can be written as

$$\Delta_1 y_{Sn+s} = -\Delta_1 y_{Sn+s-1} - \Delta_1 y_{Sn+s-2} - \cdots - \Delta_1 y_{Sn+s+1-S} + \varepsilon_{Sn+s} \tag{23}$$

and, since these autoregressive coefficients are estimated consistently, the one-step ahead forecasts are asymptotically given by $\Delta_1 \widehat{y}_{T+1|T} = -\Delta_1 y_T - \Delta_1 y_{T-1} - \cdots - \Delta_1 y_{T-S+2}$. Therefore, augmenting with $S-1$ lags of the dependent variable [see Ghysels, Lee and Siklos (1993) and Rodrigues (2000)] asymptotically implies $E[(y_{T+1} - \widehat{y}_{T+1|T})^2] = E(y_{T+1-S} + \varepsilon_{T+1} - (y_T - \Delta_1 y_T - \Delta_1 y_{T-1} - \cdots - \Delta_1 y_{T-S+2}))^2] = E[(y_{T+1-S} + \varepsilon_{T+1} - y_{T+1-S})^2] = \sigma^2$. If fewer than $S-1$ lags of the dependent variable ($\Delta_1 y_{Sn+s}$) are used, then neglected nonstationarity remains and the MSFE is anticipated to be degenerate, as in the naive case.

Turning to the SARIMA model, note that the DGP (22) can be written as

$$\Delta_1 \Delta_S y_{Sn+s} = \Delta_1 \varepsilon_{Sn+s} = \upsilon_{Sn+s} \tag{24}$$

where υ_{Sn+s} here is a noninvertible moving average process, with variance $E[(\upsilon_{Sn+s})^2] = 2\sigma^2$. Again supposing that the naive forecaster assumes υ_{Sn+s} is iid, then, using (7),

$$\begin{aligned}
&E\big[(y_{T+1} - \widehat{y}_{T+1|T})^2\big] \\
&\quad = \big[\big((y_{T+1-S} + \varepsilon_{T+1}) - (y_{T+1-S} + \Delta_S y_T + \Delta_1 \Delta_S \widehat{y}_{T+1|T})\big)^2\big] \\
&\quad = E\big[(\varepsilon_{T+1} - \Delta_S y_T)^2\big] \\
&\quad = E\big[(\varepsilon_{T+1} - \varepsilon_T)^2\big] = 2\sigma^2
\end{aligned}$$

where our naive forecaster uses $\Delta_1 \Delta_S \widehat{y}_{T+1|T} = 0$ based on iid υ_{Sn+s}. This represents an extreme case, since in practice we anticipate that some account would be taken of the autocorrelation inherent in (24). Nevertheless, it is indicative of potential forecasting problems from using an overdifferenced model, which implies the presence of noninvertible moving average unit roots that cannot be well approximated by finite order AR polynomials.

2.4.2. *Deterministic seasonal AR(1)*

Consider now a DGP of a random walk with deterministic seasonal effects, which is

$$y_{Sn+s} = y_{Sn+s-1} + \sum_{i=1}^{S} \delta_i^* D_{i,Sn+s} + \varepsilon_{Sn+s} \tag{25}$$

where $\delta_i^* = \delta_i - \delta_{i-1}$ and $\varepsilon_{Sn+s} \sim \mathrm{iid}(0, \sigma^2)$. As usual, the one-step ahead MSFE is $E[(y_{T+1} - \widehat{y}_{T+1|T})^2] = \sigma^2$ when $\widehat{y}_{T+1}$ is forecast from the correctly specified model (25), so that $\widehat{y}_{T+1|T} = y_T + \sum_{i=1}^{S} \delta_i^* D_{i,T+1}$.

If the seasonally integrated model (12) is adopted for forecasting, application of the differencing filter eliminates the deterministic seasonality and induces artificial moving average autocorrelation, since

$$\Delta_S y_{Sn+s} = \delta + S(L)\varepsilon_{Sn+s} = \delta + \upsilon_{Sn+s} \tag{26}$$

where $\delta = \sum_{i=1}^{S} \delta_i^*$, $S(L) = 1 + L + \cdots + L^{S-1}$ and here the disturbance $\upsilon_{Sn+s} = S(L)\varepsilon_{Sn+s}$ is a noninvertible moving average process, with moving average unit roots at each of the seasonal frequencies. However, even if this autocorrelation is not accounted for, δ in (26) can be consistently estimated. Although we would again expect a forecaster to recognize the presence of autocorrelation, the noninvertible moving average process cannot be approximated through the usual practice of autoregressive augmentation. Hence, as an extreme case, we again examine the consequences of a naive researcher assuming υ_{Sn+s} to be iid. Now, using the representation considered in (13) to derive the level forecast from a seasonally integrated model, it follows that

$$E\left(y_{T+1} - \widehat{y}_{T+1|T}\right)^2 = E\left[\left(y_T + \sum_{i=1}^{S} \delta_i^* D_{i,T+1} + \varepsilon_{T+1}\right) - \left(y_{T+1-S} + \Delta_S \widehat{y}_{T+1|T}\right)\right]^2$$

with $y_{T+1-S} = y_{T-S} + \sum_{i=1}^{S} \delta_i^* D_{i,T+1-S} + \varepsilon_{T+1-S}$. Note that although the seasonally integrated model apparently makes no allowance for the deterministic seasonality in the DGP, this deterministic seasonality is also present in the past observation y_{T+1-S} on which the forecast is based. Hence, since $D_{i,T+1} = D_{i,T+1-S}$, the deterministic seasonality cancels between y_T and y_{T-S}, so that

$$\begin{aligned} E\left[\left(y_{T+1} - \widehat{y}_{T+1|T}\right)^2\right] &= E\left[(y_T + \varepsilon_{T+1}) - (y_{T-S} + \varepsilon_{T+1-S})\right]^2 \\ &= E\left[(y_T - y_{T-S} - \delta + \varepsilon_{T+1} - \varepsilon_{T+1-S})^2\right] \\ &= E\left[\left((\varepsilon_T + \varepsilon_{T-1} + \cdots + \varepsilon_{T-S+1}) + \varepsilon_{T+1} - \varepsilon_{T+1-S}\right)^2\right] \\ &= E\left[(\varepsilon_{T+1} + \varepsilon_T + \cdots + \varepsilon_{T-S+2})^2\right] = S\sigma^2 \end{aligned}$$

as, from (26), the naive forecaster uses $\Delta_S \widehat{y}_{T+1} = \delta$. The result also uses (26) to substitute for $y_T - y_{T-S}$. Thus, as a consequence of seasonal overdifferencing, the MSFE increases proportionally to the periodicity of the data. This MSFE effect can, however, be reduced if the overdifferencing is (partially) accounted for through augmentation.

Now consider the use of the SARIMA model when the data is in fact generated by (25). Although

$$\Delta_1 \Delta_S y_{Sn+s} = \Delta_S \varepsilon_{Sn+s} \tag{27}$$

we again consider the naive forecaster who assumes $\upsilon_{Sn+s} = \Delta_S \varepsilon_{Sn+s}$ is iid. Using (7), and noting from (27) that the forecaster uses $\Delta_1 \Delta_S \widehat{y}_{T+1} = 0$, it follows that

$$\begin{aligned} E\left[\left(y_{T+1} - \widehat{y}_{T+1|T}\right)^2\right] &= E\left[\left(y_T + \sum_{i=1}^{S} \delta_i^* D_{i,T+1} + \varepsilon_{T+1} - y_{T+1-S} + \Delta_S y_T\right)^2\right] \\ &= E\left[(\varepsilon_{T+1} - \varepsilon_{T+1-S})^2\right] = 2\sigma^2. \end{aligned}$$

Once again, the deterministic seasonal pattern is taken into account indirectly, through the implicit dependence of the forecast on the past observed value y_{T+1-S} that incorporates the deterministic seasonal effects. Curiously, although the degree of overdifferencing is higher in the SARIMA than in the seasonally integrated model, the MSFE is smaller in the former case.

As already noted, our analysis here does not take account of either augmentation or parameter estimation and hence these results or misspecified models may be considered "worst case" scenarios. It is also worth noting that when seasonally integrated or SARIMA models are used for forecasting a deterministic seasonality DGP, then fewer parameters might be estimated in practice than required in the true DGP. This greater parsimony may outweigh the advantages of using the correct specification and hence it is plausible that a misspecified model could, in particular cases and in moderate or small samples, yield lower MSFE. These issues are investigated in the next subsection through a Monte Carlo analysis.

2.4.3. Monte Carlo analysis

This Monte Carlo analysis complements the results of the previous subsection, allowing for augmentation and estimation uncertainty. In all experiments, 10000 replications are used with a maximum lag order considered of $p\max = 8$, the lag selection based on Ng and Perron (1995). Forecasts are performed for horizons $h = 1, \ldots, 8$, in samples of $T = 100$, 200 and 400 observations. The tables below report results for $h = 1$ and $h = 8$.

Forecasts are generated using the following three types of models:

$$M_1\colon \quad \Delta_1 \Delta_4 y_{4n+s} = \sum_{i=1}^{p_1} \phi_{1,i} \Delta_1 \Delta_4 y_{4n+s-i} + \varepsilon_{1,4n+s},$$

$$M_2\colon \quad \Delta_4 y_{4n+s} = \sum_{i=1}^{p_2} \phi_{2,i} \Delta_4 y_{4n+s-i} + \varepsilon_{2,4n+s},$$

$$M_3\colon \quad \Delta_1 y_{4n+s} = \sum_{k=1}^{4} \delta_k D_{k,4n+s} + \sum_{i=1}^{p_3} \phi_{3,i} \Delta_1 y_{4n+s-i} + \varepsilon_{3,4n+s}.$$

The first DGP is the seasonal autoregressive process

$$y_{Sn+s} = \rho y_{S(n-1)+s} + \varepsilon_{Sn+s} \tag{28}$$

where $\varepsilon_{Sn+s} \sim \text{niid}(0, 1)$ and $\rho = \{1, 0.9, 0.8\}$.

Panels (a) to (c) of Table 1 indicate that as one moves from $\rho = 1$ into the stationarity region ($\rho = 0.9$, $\rho = 0.8$) the one-step ahead ($h = 1$) empirical MSFE deteriorates for all forecasting models. For $h = 8$, a similar phenomenon occurs for M_1 and M_2, however M_3 shows some improvement. This behavior is presumably related to the greater degree of overdifferencing imposed by models M_1 and M_2, compared to M_3.

When $\rho = 1$, panel (a) indicates that model M_2 (which considers the correct degree of differencing) yields lower MSFE for both $h = 1$ and $h = 8$ than M_1 and M_3. This advantage for M_2 carries over in relation to M_1 even when $\rho < 1$. However, in panel (c),

Table 1
MSFE when the DGP is (28)

h	T	(a) $\rho = 1$ M_1	M_2	M_3	(b) $\rho = 0.9$ M_1	M_2	M_3	(c) $\rho = 0.8$ M_1	M_2	M_3
1	100	1.270	1.035	1.136	1.347	1.091	1.165	1.420	1.156	1.174
	200	1.182	1.014	1.057	1.254	1.068	1.074	1.324	1.123	1.087
	400	1.150	1.020	1.041	1.225	1.074	1.044	1.294	1.123	1.058
8	100	2.019	1.530	1.737	2.113	1.554	1.682	2.189	1.579	1.585
	200	1.933	1.528	1.637	2.016	1.551	1.562	2.084	1.564	1.483
	400	1.858	1.504	1.554	1.942	1.533	1.485	2.006	1.537	1.421
					Average number of lags					
	100	5.79	1.21	3.64	5.76	1.25	3.65	5.81	1.39	3.71
	200	6.98	1.21	3.64	6.94	1.30	3.67	6.95	1.57	3.79
	400	7.65	1.21	3.62	7.67	1.38	3.70	7.68	1.88	3.97

as one moves further into the stationarity region ($\rho = 0.8$) the performance of M_3 is superior to M_2 for sample sizes $T = 200$ and $T = 400$.

Our simple analysis of the previous subsection shows that M_3 should (asymptotically and with augmentation) yield the same forecasts as M_2 for the seasonal random walk of panel (a), but less accurate forecasts are anticipated from M_1 in this case. Our Monte Carlo results verify the practical impact of that analysis. Interestingly, the autoregressive order selected remains relatively stable across the three autoregressive scenarios considered ($\rho = 1, 0.9, 0.8$). Indeed, in this and other respects, the "close to nonstationary" DGPs have similar forecast implications as the nonstationary random walk.

The second DGP considered in this simulation is the first order autoregressive process with deterministic seasonality,

$$y_{Sn+s} = \sum_{i=1}^{S} \delta_i D_{i,Sn+s} + x_{Sn+s}, \tag{29}$$

$$x_{Sn+s} = \rho x_{Sn+s-1} + \varepsilon_{Sn+s} \tag{30}$$

where $\varepsilon_{Sn+s} \sim \text{niid}(0, 1)$, $\rho = \{1, 0.9, 0.8\}$ and $(\delta_1, \delta_2, \delta_3, \delta_4) = (-1, 1, -1, 1)$. Here M_3 provides the correct DGP when $\rho = 1$.

Table 2 shows that (as anticipated) M_3 outperforms M_1 and M_2 when $\rho = 1$, and this carries over to $\rho = 0.9, 0.8$ when $h = 1$. It is also unsurprising that M_3 yields lowest MSFE for $h = 8$ when this is the true DGP in panel (a). Although our previous analysis indicates that M_2 should perform worse than M_1 in this case when the models are not augmented, in practice these models have similar performance when $h = 1$ and M_2 is superior at $h = 8$. The superiority of M_3 also applies when $\rho = 0.9$. However, despite greater overdifferencing, M_2 outperforms M_3 at $h = 8$ when $\rho = 0.8$. In this case, the estimation of additional parameters in M_3 appears to have an adverse effect on forecast

Table 2
MSFE when the DGP is (29) and (30)

		(a) $\rho = 1$			(b) $\rho = 0.9$			(c) $\rho = 0.8$		
h	T	M_1	M_2	M_3	M_1	M_2	M_3	M_1	M_2	M_3
1	100	1.426	1.445	1.084	1.542	1.472	1.151	1.626	1.488	1.210
	200	1.370	1.357	1.032	1.478	1.387	1.092	1.550	1.401	1.145
	400	1.371	1.378	1.030	1.472	1.402	1.077	1.538	1.416	1.120
8	100	7.106	5.354	4.864	6.831	4.073	3.993	5.907	3.121	3.246
	200	7.138	5.078	4.726	6.854	3.926	3.887	5.864	3.030	3.139
	400	7.064	4.910	4.577	6.774	3.839	3.771	5.785	2.986	3.003
					Average number of lags					
	100	2.64	4.07	0.80	2.68	4.22	1.00	2.86	4.27	1.48
	200	2.70	4.34	0.78	2.76	4.46	1.24	3.16	4.49	2.36
	400	2.71	4.48	0.76	2.81	4.53	1.72	3.62	4.53	4.02

accuracy, compared with M_2. In this context, note that the number of lags used in M_3 is increasing as one moves into the stationarity region.

One striking finding of the results in Tables 1 and 2 is that M_2 and M_3 have similar forecast performance at the longer forecast horizon of $h = 8$, or two years. In this sense, the specification of seasonality as being of the nonstationary stochastic or deterministic form may not be of great concern when forecasting. However, the two zero frequency unit roots imposed by the SARIMA model M_1 (and not present in the DGP) leads to forecasts at this non-seasonal horizon which are substantially worse than those of the other two models.

At one-step-ahead horizon, if it is unclear whether the process has zero and seasonal unit roots, our results indicate that the use of the deterministic seasonality model with augmentation may be a more flexible tool than the seasonally integrated model.

2.5. Seasonal cointegration

The univariate models addressed in the earlier subsections are often adequate when short-run forecasts are required. However, multivariate models allow additional information to be utilized and may be expected to improve forecast accuracy. In the context of nonstationary economic variables, cointegration restrictions can be particularly important. There is a vast literature on the forecasting performance of cointegrated models, including Ahn and Reinsel (1994), Clements and Hendry (1993), Lin and Tsay (1996) and Christoffersen and Diebold (1998). The last of these, in particular, shows that the incorporation of cointegration restrictions generally leads to improved long-run forecasts.

Despite the vast literature concerning cointegration, that relating specifically to the seasonal context is very limited. This is partly explained by the lack of evidence for the presence of the full set of seasonal unit roots in economic time series. If seasonality is of the deterministic form, with nonstationarity confined to the zero frequency, then conventional cointegration analysis is applicable, provided that seasonal dummy variables are included where appropriate. Nevertheless, seasonal differencing is sometimes required and it is important to investigate whether cointegration applies also to the seasonal frequency, as well as to the conventional long-run (at the zero frequency). When seasonal cointegration applies, we again anticipate that the use of these restrictions should improve forecast performance.

2.5.1. *Notion of seasonal cointegration*

To introduce the concept, now let y_{Sn+s} be a vector of seasonally integrated time series. For expositional purposes, consider the quarterly ($S = 4$) case

$$\Delta_4 y_{4n+s} = \eta_{4n+s} \tag{31}$$

where η_{4n+s} is a zero mean stationary and invertible vector stochastic process. Given the vector of seasonally integrated time series, linear combinations may exist that cancel out corresponding seasonal (as well as zero frequency) unit roots. The concept of seasonal cointegration is formalized by Engle, Granger and Hallman (1989), Hylleberg, Engle, Granger and Yoo [HEGY] (1990) and Engle et al. (1993). Based on HEGY, the error-correction representation of a quarterly seasonally cointegrated vector is[4]

$$\begin{aligned}\beta(L)\Delta_4 y_{4n+s} = {}& \alpha_0 b_0' y_{0,4n+s-1} + \alpha_{11} b_{11}' y_{1,4n+s-1} + \alpha_{12} b_{12}' y_{1,4n+s-2} \\ & + \alpha_2 b_2' y_{2,4n+s-1} + \varepsilon_{4n+s}\end{aligned} \tag{32}$$

where ε_{4n+s} is an iid process, with covariance matrix $E[\varepsilon_{4n+s}\varepsilon_{4n+s}'] = \Sigma$ and each element of the vector $y_{i,4n+s}$ ($i = 0, 1, 2$) is defined through the transformations of (11). Since each element of y_{4n+s} exhibits nonstationarity at the zero and the two seasonal frequencies (π, $\pi/2$), cointegration may apply at each of these frequencies. Indeed, in general, the rank as well as the coefficients of the cointegrating vectors may differ over these frequencies.

The matrix b_0 of (32) contains the linear combinations that eliminate the zero frequency unit root (+1) from the individual $I(1)$ series of $y_{0,4n+s}$. Similarly, b_2 cancels the Nyquist frequency unit root (−1), i.e., the nonstationary biannual cycle present in $y_{2,4n+s}$. The coefficient matrices α_0 and α_2 represent the adjustment coefficients for the variables of the system to the cointegrating relationships at the zero and biannual frequencies, respectively. For the annual cycle corresponding to the complex pair of unit roots $\pm i$, the situation is more complex, leading to two terms in (32). The fact that the cointegrating relations (b_{12}', b_{11}') and adjustment matrices (α_{12}, α_{11}) relate to two lags of $y_{1,4n+s}$ is called polynomial cointegration by Lee (1992).

Residual-based tests for the null hypothesis of no seasonal cointegration are discussed by Engle et al. (1993) in the setup of single equation regression models, while Hassler

[4] The generalization for seasonality at any frequency is discussed in Johansen and Schaumburg (1999).

and Rodrigues (2004) provide an empirically more appealing approach. Lee (1992) developed the first system approach to testing for seasonal cointegration, extending the analysis of Johansen (1988) to this case. However, Lee assumes $\alpha_{11}b'_{11} = 0$, which Johansen and Schaumburg (1999) argue is restrictive and they provide a more general treatment.

2.5.2. Cointegration and seasonal cointegration

Other representations may shed light on issues associated with forecasting and seasonal cointegration. Using definitions (11), (32) can be rewritten as

$$\begin{aligned} \beta(L)\Delta_4 y_{4n+s} = {} & \Pi_1 y_{4n+s-1} + \Pi_2 y_{4n+s-2} + \Pi_3 y_{4n+s-3} \\ & + \Pi_4 y_{4(n-1)+s} + \varepsilon_{4n+s} \end{aligned} \tag{33}$$

where the matrices Π_i $(i = 1, 2, 3, 4)$ are given by

$$\begin{aligned} \Pi_1 &= \alpha_0 b'_0 - \alpha_2 b'_2 - \alpha_{11} b'_{11}, \quad & \Pi_2 &= \alpha_0 b'_0 + \alpha_2 b'_2 - \alpha_{12} b'_{12}, \\ \Pi_3 &= \alpha_0 b'_0 - \alpha_2 b'_2 + \alpha_{11} b'_{11}, \quad & \Pi_4 &= \alpha_0 b'_0 + \alpha_2 b'_2 + \alpha_{12} b'_{12}. \end{aligned} \tag{34}$$

Thus, seasonal cointegration implies that the annual change adjusts to y_{4n+s-i} at lags $i = 1, 2, 3, 4$, with (in general) distinct coefficient matrices at each lag; see also Osborn (1993).

Since seasonal cointegration is considered relatively infrequently, it is natural to ask what are the implications of undertaking a conventional cointegration analysis in the presence of seasonal cointegration. From (33) we can write, assuming $\beta(L) = 1$ for simplicity, that,

$$\begin{aligned} \Delta_1 y_{4n+s} &= (\Pi_1 - I)y_{4n+s-1} + \Pi_2 y_{4n+s-2} + \Pi_3 y_{4n+s-3} \\ &\quad + (\Pi_4 + I)y_{4n+s-4} + \varepsilon_{4n+s} \\ &= (\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4)y_{4n+s-1} - (\Pi_2 + \Pi_3 + \Pi_4 + I)\Delta_1 y_{4n+s-1} \\ &\quad - (\Pi_3 + \Pi_4 + I)\Delta_1 y_{4n+s-2} + (\Pi_4 + I)\Delta_1 y_{4n+s-3} + \varepsilon_{4n+s}. \end{aligned} \tag{35}$$

Thus (provided that the ECM is adequately augmented with at least three lags of the vector of first differences), a conventional cointegration analysis implies (35), where the matrix coefficient on the lagged level y_{4n+s-1} is $\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4$. However, it is easy to see from (34) that

$$\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4 = 4\alpha_0 b'_0, \tag{36}$$

so that a conventional cointegration analysis should uncover the zero frequency cointegrating relationships. Although the cointegrating relationships at seasonal frequencies do not explicitly enter the cointegration considered in (36), these will be reflected in the coefficients for the lagged first difference variables, as implied by (35). This generalizes the univariate result of Ghysels, Lee and Noh (1994), that a conventional Dickey–Fuller test remains applicable in the context of seasonal unit roots, provided that the test regression is sufficiently augmented.

2.5.3. Forecasting with seasonal cointegration models

The handling of deterministic components in seasonal cointegration is discussed by Franses and Kunst (1999). In particular, the seasonal dummy variable coefficients need to be restricted to the (seasonal) cointegrating space if seasonal trends are not to be induced in the forecast series.

However, to focus on seasonal cointegration, we continue to ignore deterministic terms. The optimal forecast in a seasonally cointegrated system can then be obtained from (33) as

$$\begin{aligned}\Delta_4\widehat{y}_{T+h|T} &= \Pi_1\widehat{y}_{T+h-1|T} + \Pi_2\widehat{y}_{T+h-2|T} + \Pi_3\widehat{y}_{T+h-3|T} + \Pi_4\widehat{y}_{T+h-4|T} \\ &\quad + \sum_{i=1}^{p} \beta_i \Delta_4\widehat{y}_{T+h-i|T} \end{aligned} \tag{37}$$

where, analogously to the univariate case, $\widehat{y}_{T+h|T} = E[y_{T+h}|y_1, \ldots, y_T] = \widehat{y}_{T+h-4|T} + \Delta_4\widehat{y}_{T+h|T}$ is computed recursively for $h = 1, 2, \ldots$. As this is a linear system, optimal forecasts of another linear transformation, such as $\Delta_1\widehat{y}_{T+h}$, are obtained by applying the required linear transformation to the forecasts generated by (37).

For one-step ahead forecasts ($h = 1$), it is straightforward to see that the matrix MSFE for this system is

$$E\big[\big(y_{T+1} - \widehat{y}_{T+1|T}\big)\big(y_{T+1} - \widehat{y}_{T+1|T}\big)'\big] = E\big[\varepsilon_{T+1}\varepsilon'_{T+1}\big] = \Sigma.$$

To consider longer horizons, we take the case of $h = 2$ and assume $\beta(L) = 1$ for simplicity. Forecasting from the seasonally cointegrated system then implies

$$\begin{aligned} &E\big[\big(y_{T+2} - \widehat{y}_{T+2|T}\big)\big(y_{T+2} - \widehat{y}_{T+2|T}\big)'\big] \\ &\quad = E\big[\big\{\Pi_1\big(y_{T+1} - \widehat{y}_{T+1|T}\big) + \varepsilon_{T+2}\big\}\big\{\Pi_1\big(y_{T+1} - \widehat{y}_{T+1|T}\big) + \varepsilon_{T+2}\big\}'\big] \\ &\quad = \Pi_1\Sigma\Pi_1' + \Sigma \end{aligned} \tag{38}$$

with $\Pi_1 = (\alpha_0 b_0' - \alpha_{11}b_{11}' - \alpha_2 b_2')$. Therefore, cointegration at the seasonal frequencies plays a role here, in addition to cointegration at the zero frequency.

If the conventional ECM representation (35) is used, then (allowing for the augmentation required even when $\beta(L) = 1$) identical expressions to those just obtained result for the matrix MSFE, due to the equivalence established above between the seasonal and the conventional ECM representations.

When forecasting seasonal time series, and following the seminal paper of Davidson et al. (1978), a common approach is to model the annual differences with cointegration applied at the annual lag. Such a model is

$$\beta^*(L)\Delta_4 y_{4n+s} = \Pi y_{4(n-1)+s} + v_{4n+s} \tag{39}$$

where $\beta^*(L)$ is a polynomial in L and v_{4n+s} is assumed to be vector white noise. If the DGP is given by the seasonally cointegrated model, rearranging (23) yields

$$\beta(L)\Delta_4 y_{4n+s} = (\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4)y_{4(n-1)+s} + \Pi_1\Delta_1 y_{4n+s-1}$$

$$+ (\Pi_1 + \Pi_2)\Delta_1 y_{4n+s-2} + (\Pi_1 + \Pi_2 + \Pi_3)\Delta_1 y_{4n+s-3} + \varepsilon_{4n+s}. \quad (40)$$

As with conventional cointegration modelling in first differences, the long run zero frequency cointegrating relationships may be uncovered by such an analysis, through $\Pi_1 + \Pi_2 + \Pi_3 + \Pi_4 = \Pi = 4\alpha_0 b_0'$. However, the autoregressive augmentation in $\Delta_4 y_{4n+s}$ adopted in (39) implies overdifferencing compared with the first difference terms on the right-hand side of (40), and hence is unlikely (in general) to provide a good approximation to the coefficients of $\Delta_1 y_{4n+s-i}$ of (40). Indeed, the model based on (39) is valid only when $\Pi_1 = \Pi_2 = \Pi_3 = 0$.

Therefore, if a researcher wishes to avoid issues concerned with seasonal cointegration when such cointegration may be present, it is preferable to use a conventional VECM (with sufficient augmentation) than to consider an annual difference specification such as (39).

2.5.4. *Forecast comparisons*

Few papers examine forecasts for seasonally cointegrated models for observed economic time series against the obvious competitors of conventional vector error-correction models and VAR models in first differences. In one such comparison, Kunst (1993) finds that accounting for seasonal cointegration generally provides limited improvements, whereas Reimers (1997) finds seasonal cointegration models produce relatively more accurate forecasts when longer forecast horizons are considered. Kunst and Franses (1998) show that restricting seasonal dummies in seasonal cointegration yields better forecasts in most cases they consider, which is confirmed by Löf and Lyhagen (2002). From a Monte Carlo study, Lyhagen and Löf (2003) conclude that use of the seasonal cointegration model provides a more robust forecast performance than models based on pre-testing for unit roots at the zero and seasonal frequencies.

Our review above of cointegration and seasonal cointegration suggests that, in the presence of seasonal cointegration, conventional cointegration modelling will uncover zero frequency cointegration. Since seasonality is essentially an intra-year phenomenon, it may be anticipated that zero frequency cointegration may be relatively more important than seasonal cointegration at longer forecast horizons. This may explain the findings of Kunst (1993) and Reimers (1997) that conventional cointegration models often forecast relatively well in comparison with seasonal cointegration. Our analysis also suggests that a model based on (40) should not, in general, be used for forecasting, since it does not allow for the possible presence of cointegration at the seasonal frequencies.

2.6. *Merging short- and long-run forecasts*

In many practical contexts, distinct models are used to generate forecasts at long and short horizons. Indeed, long-run models may incorporate factors such as technical progress, which are largely irrelevant when forecasting at a horizon of (say) less than a year. In an interesting paper Engle, Granger and Hallman (1989) discuss merging short- and long-run forecasting models. They suggest that when considering a (single) variable y_{Sn+s}, one can think of models generating the short- and long-run forecasts as approximating different parts of the DGP, and hence these models may have different specifications with non-overlapping sets of explanatory variables. For instance, if y_{Sn+s} is monthly demand for electricity (as considered by Engle, Granger and Hallman), the short-run model may concentrate on rapidly changing variables, including strongly seasonal ones (e.g., temperature and weather variables), whereas the long-run model assimilates slowly moving variables, such as population characteristics, appliance stock and efficiencies or local output. To employ all the variables in the short-run model is too complex and the long-run explanatory variables may not be significant when estimation is by minimization of the one-month forecast variance.

Following Engle, Granger and Hallman (1989), consider $y_{Sn+s} \sim I(1)$ which is cointegrated with variables of the $I(1)$ vector x_{Sn+s} such that $z_{Sn+s} = y_{Sn+s} - \alpha_1' x_{Sn+s}$ is stationary. The true DGP is

$$\Delta_1 y_{Sn+s} = \delta - \gamma z_{Sn+s-1} + \beta' w_{Sn+s} + \varepsilon_{Sn+s}, \tag{41}$$

where w_{Sn+s} is a vector of $I(0)$ variables that can include lags of $\Delta_1 y_{Sn+s}$. Three forecasting models can be considered: the complete true model given by (41), the long-run forecasting model of $y_{Sn+s} = \alpha_0 + \alpha_1' x_{Sn+s} + \eta_{Sn+s}$ and the short-run forecasting model that omits the error-correction term z_{Sn+s-1}. For convenience, we assume that annual forecasts are produced from the long-run model, while forecasts of seasonal (e.g., monthly or quarterly) values are produced by the short-run model.

If all data are available at a seasonal periodicity and the DGP is known, one-step forecasts can be found using (41) as

$$\widehat{y}_{T+1|T} = \delta - (1+\gamma) y_T + \gamma \alpha_1' x_T + \beta' \widehat{w}_{T+1|T}. \tag{42}$$

Given forecasts of x and w, multi-step forecasts $\widehat{y}_{T+h|T}$ can be obtained by iterating (42) to the required horizon. For forecasting a particular season, the long-run forecasts of w_{Sn+s} are constants (their mean for that season) and the DGP implies the long-run forecast

$$\widehat{y}_{T+h|T} \approx \alpha_1' \widehat{x}_{T+h|T} + c \tag{43}$$

where c is a (seasonally varying) constant. Annual forecasts from (43) will be produced by aggregating over seasons, which removes seasonal effects in c. Consequently, the long-run forecasting model should produce annual forecasts similar to those from (43) using the DGP. Similarly, although the short-run forecasting model omits the error-correction term z_{Sn+s}, it will be anticipated to produce similar forecasts to (42), since season-to-season fluctuations will dominate short-run forecasts.

Due to the unlikely availability of long-run data at the seasonal frequency, the complete model (41) is unattainable in practice. Essentially, Engle, Granger and Hallman (1989) propose that the forecasts from the long-run and short-run models be combined to produce an approximation to this DGP. Although not discussed in detail by Engle, Granger and Hallman (1989), long-run forecasts may be made at the annual frequency and then interpolated to seasonal values, in order to provide forecasts approximating those from (41).

In this set-up, the long-run model includes annual variables and has nothing to say about seasonality. By design, cointegration relates only to the zero frequency. Seasonality is allocated entirely to the short-run and is modelled through the deterministic component and the forecasts $\widehat{w}_{T+h|T}$ of the stationary variables. Rather surprisingly, this approach to forecasting appears almost entirely unexplored in subsequent literature, with issues of seasonal cointegration playing a more prominent role. This is unfortunate, since (as noted in the previous subsection) there is little evidence that seasonal cointegration improves forecast accuracy and, in any case, can be allowed for by including sufficient lags of the relevant variables in the dynamics of the model. In contrast, the approach of Engle, Granger and Hallman (1989) allows information available only at an annual frequency to play a role in capturing the long-run, and such information is not considered when the researcher focuses on seasonal cointegration.

3. Periodic models

Periodic models provide another approach to modelling and forecasting seasonal time series. These models are more general than those discussed in the previous section in allowing all parameters to vary across the seasons of a year. Periodic models can be useful in capturing economic situations where agents show distinct seasonal characteristics, such as seasonally varying utility of consumption [Osborn (1988)]. Within economics, periodic models usually take an autoregressive form and are known as PAR (periodic autoregressive) models.

Important developments in this field, have been made by, *inter alia*, Pagano (1978), Troutman (1979), Gladyshev (1961), Osborn (1991), Franses (1994) and Boswijk and Franses (1996). Applications of PAR models include, for example, Birchenhall et al. (1989), Novales and Flores de Fruto (1997), Franses and Romijn (1993), Herwartz (1997), Osborn and Smith (1989) and Wells (1997).

3.1. Overview of PAR models

A univariate PAR(p) model can be written as

$$y_{Sn+s} = \sum_{j=1}^{S} \bigl[\mu_j + \tau_j(Sn+s)\bigr] D_{j,Sn+s} + x_{Sn+s}, \tag{44}$$

$$x_{Sn+s} = \sum_{j=1}^{S} \sum_{i=1}^{p_j} \phi_{ij} D_{j,Sn+s} x_{Sn+s-i} + \varepsilon_{Sn+s} \tag{45}$$

where (as in the previous section) S represents the periodicity of the data, while here p_j is the order of the autoregressive component for season j, $p = \max(p_1, \ldots, p_S)$, $D_{j,Sn+s}$ is again a seasonal dummy that is equal to 1 in season j and zero otherwise, and $\varepsilon_{Sn+s} \sim \text{iid}(0, \sigma_s^2)$. The PAR model of (44)–(45) requires a total of $(3S + \sum_{j=1}^{S} p_j)$ parameters to be estimated. This basic model can be extended by including periodic moving average terms [Tiao and Grupe (1980), Lütkepohl (1991)].

Note that this process is nonstationary in the sense that the variances and covariances are time-varying within the year. However, considered as a vector process over the S seasons, stationarity implies that these intra-year variances and covariances remain constant over years, $n = 0, 1, 2, \ldots$. It is this vector stationarity concept that is appropriate for PAR processes.

Substituting from (45) into (44), the model for season s is

$$\phi_s(L) y_{Sn+s} = \phi_s(L)\big[\mu_s + \tau_s(Sn + s)\big] + \varepsilon_{Sn+s} \tag{46}$$

where $\phi_j(L) = 1 - \phi_{1j} L - \cdots - \phi_{p_j,j} L^{p_j}$. Alternatively, following Boswijk and Franses (1996), the model for season s can be represented as

$$(1 - \alpha_s L) y_{Sn+s} = \delta_s + \omega_s(Sn + s) + \sum_{k=1}^{p-1} \beta_{ks}(1 - \alpha_{s-k} L) y_{Sn+s-k} + \varepsilon_{Sn+s} \tag{47}$$

where $\alpha_{s-Sm} = \alpha_s$ for $s = 1, \ldots, S$, $m = 1, 2, \ldots$ and $\beta_j(L)$ is a $(p_j - 1)$-order polynomial in L. Although the parameterization of (47) is useful, it should also be appreciated that the factorization of $\phi_s(L)$ implied in (47) is not, in general, unique [del Barrio Castro and Osborn (2004)]. Nevertheless, this parameterization is useful when the unit root properties of y_{Sn+s} are isolated in $(1 - \alpha_s L)$. In particular, the process is said to be periodically integrated if

$$\prod_{s=1}^{S} \alpha_s = 1, \tag{48}$$

with the stochastic part of $(1 - \alpha_s L) y_{Sn+s}$ being stationary. In this case, (48) serves to identify the parameters of (47) and the model is referred to as a periodic integrated autoregressive (PIAR) model. To distinguish periodic integration from conventional (nonperiodic) integration, we require that not all $\alpha_s = 1$ in (48).

An important consequence of periodic integration is that such series cannot be decomposed into distinct seasonal and trend components; see Franses (1996, Chapter 8). An alternative possibility to the PIAR process is a conventional unit root process with periodic stationary dynamics, such as

$$\beta_s(L) \Delta_1 y_{Sn+s} = \delta_s + \varepsilon_{Sn+s}. \tag{49}$$

As discussed below, (47) and (49) have quite different forecast implications for the future pattern of the trend.

3.2. Modelling procedure

The crucial issues for modelling a potentially periodic process are deciding whether the process is, indeed, periodic and deciding the appropriate order p for the PAR.

3.2.1. Testing for periodic variation and unit roots

Two approaches can be considered to the inter-related issues of testing for the presence of periodic coefficient variation.

(a) Test the nonperiodic (constant autoregressive coefficient) null hypothesis

$$H_0\colon \quad \phi_{ij} = \phi_i, \quad j = 1, \dots, S, i = 1, \dots, p \tag{50}$$

against the alternative of a periodic model using a χ^2 or F test (the latter might be preferred unless the number of years of data is large). This is conducted using an OLS estimation of (44) and, as no unit root restriction is involved, its validity does not depend on stationarity [Boswijk and Franses (1996)].

(b) Estimate a nonperiodic model and apply a diagnostic test for periodic autocorrelation to the residuals [Franses (1996, pp. 101–102)]. Further, Franses (1996) argues that neglected parameter variations may surface in the variance of the residual process, so that a test for periodic heteroskedasticity can be considered, by regressing the squared residuals on seasonal dummy variables [see also del Barrio Castro and Osborn (2004)]. These can again be conducted using conventional distributions.

Following a test for periodic coefficient variation, such as (50), unit root properties may be examined. Boswijk and Franses (1996) develop a generalization of the Dickey–Fuller unit root t-test statistic applicable in a periodic context. Conditional on the presence of a unit root, they also discuss testing the restriction $\alpha_s = 1$ in (47), with this latter test being a test of restrictions that can be applied using the conventional χ^2 or F-distribution. When the restrictions $\alpha_s = 1$ are valid, the process can be written as (49) above. Ghysels, Hall and Lee (1996) also propose a test for seasonal integration in the context of a periodic process.

3.2.2. Order selection

The order selection of the autoregressive component of the PAR model is obviously important. Indeed, because the number of autoregressive coefficients required is (in general) pS, this may be considered to be more crucial in this context than for the linear AR models of the previous section.

Order specification is frequently based on an information criterion. Franses and Paap (1994) find that the Schwarz Information Criterion (SIC) performs better for order selection in periodic models than the Akaike Information Criterion (AIC). This is, perhaps, unsurprising in that AIC leads to more highly parameterized models, which may be considered overparameterized in the periodic context. Franses and Paap (1994) recommend backing up the SIC strategy that selects p by F-tests for $\phi_{i,p+1} = 0, i = 1, \dots, S$. Having established the PAR order, the null hypothesis of nonperiodicity (50) is then examined.

If used without restrictions, a PAR model tends to be highly parameterized, and the application of restrictions may yield improved forecast accuracy. Some of the model reduction strategies that can be considered are:

- Allow different autoregressive orders p_j for each season, $j = 1, \dots, S$, with possible follow-up elimination of intermediate regressors by an information criterion or using statistical significance.
- Employ common parameters for across seasons. Rodrigues and Gouveia (2004) specify a PAR model for monthly data based on $S = 3$ seasons. In the same vein, Novales and Flores de Fruto (1997) propose grouping similar seasons into blocks to reduce the number of periodic parameters to be estimated.
- Reduce the number of parameters by using short Fourier series [Jones and Brelsford (1968), Lund et al. (1995)]. Such Fourier reductions are particularly useful when changes in the correlation structure over seasons are not abrupt.
- Use a layered approach, where a "first layer" removes the periodic autocorrelation in the series, while a "second layer" has an ARMA(p, q) representation [Bloomfield, Hurd and Lund (1994)].

3.3. Forecasting with univariate PAR models

Perhaps the simplest representation of a PAR model for forecasting purposes is (47), from which the h-step forecast is given by

$$\begin{aligned}\widehat{y}_{T+h|T} &= \alpha_s \widehat{y}_{T+h-1|T} + \delta_s + \omega_s(T+h)\\ &\quad + \sum_{k=1}^{p-1} \beta_{ks}\big(\widehat{y}_{T+h-k|T} - \alpha_{s-k}\widehat{y}_{T+h-k-1|T}\big)\end{aligned} \tag{51}$$

when $T+h$ falls in season s. This expression can be iterated for $h = 1, 2, \dots$. Assuming a unit root PAR process, we can distinguish the forecasting implications of y being periodically integrated (with $\prod_{i=1}^{S} \alpha_i = 1$, but not all $\alpha_s = 1$) and an $I(1)$ process ($\alpha_s = 1, s = 1, \dots, S$).

To discuss the essential features of the $I(1)$ case, an order $p = 2$ is sufficient. A key feature for forecasting nonstationary processes is the implications for the deterministic component. In this specific case, $\phi_s(L) = (1-L)(1-\beta_s L)$, so that (46) and (47) imply

$$\begin{aligned}\delta_s + \omega_s(T+h) &= (1-L)(1-\beta_s L)\big[\mu_s + \tau_s(T+h)\big]\\ &= \Delta\mu_s - \beta_s\Delta\mu_{s-1} + \tau_s(T+h) - (1+\beta_s)\tau_{s-1}(T+h-1)\\ &\quad + \beta_s\tau_{s-2}(T+h-2)\end{aligned}$$

and hence

$$\delta_s = \Delta\mu_s - \beta_s\Delta\mu_{s-1} + \tau_{s-1} + \beta_s\tau_{s-1} - 2\beta_s\tau_{s-2},$$

$$\omega_s = \tau_s - (1 + \beta_s)\tau_{s-1} + \beta_s \tau_{s-2}.$$

Excluding specific cases of interaction[5] between values of τ_s and β_s, the restriction $\omega_s = 0, s = 1, \ldots, S$ in (51) implies $\tau_s = \tau$, so that the forecasts for the seasons do not diverge as the forecast horizon increases. With this restriction, the intercept

$$\delta_s = \Delta\mu_s - \beta_s \Delta\mu_{s-1} + (1 - \beta_s)\tau$$

implies a deterministic seasonal pattern in the forecasts. Indeed, in the special case that $\beta_s = \beta, s = 1, \ldots, S$, this becomes the forecast for a deterministic seasonal process with a stationary AR(1) component.

The above discussion shows that a stationary periodic autoregression in an $I(1)$ process does not essentially alter the characteristics of the forecasts, compared with an $I(1)$ process with deterministic seasonality. We now turn attention to the case of periodic integration.

In a PIAR process, the important feature is the periodic nonstationarity, and hence we gain sufficient generality for our discussion by considering $\phi_s(L) = 1 - \alpha_s L$. In this case, (51) becomes

$$\widehat{y}_{T+h|T} = \alpha_s \widehat{y}_{T+h-1|T} + \delta_s + \omega_s(T + h) \tag{52}$$

for which (46) implies

$$\begin{aligned} \delta_s + \omega_s(T + h) &= (1 - \alpha_s L)\bigl[\mu_s + \tau_s(T + h)\bigr] \\ &= \mu_s - \alpha_s \mu_{s-1} + \tau_s(T + h) - \alpha_s \tau_{s-1}(T + h - 1) \end{aligned}$$

and hence

$$\delta_s = \mu_s - \alpha_s \mu_{s-1} + \alpha_s \tau_{s-1},$$

$$\omega_s = \tau_s - \alpha_s \tau_{s-1}.$$

Here imposition of $\omega_s = 0$ $(s = 1, \ldots, S)$ implies $\tau_s - \alpha_s \tau_{s-1} = 0$, and hence $\tau_s \neq \tau_{s-1}$ in (44) for at least one s, since the periodic integrated process requires not all $\alpha_s = 1$. Therefore, forecasts exhibiting distinct trends over the S seasons are a natural consequence of a PIAR specification, whether or not an explicit trend is included in (52). A forecaster adopting a PIAR model needs to appreciate this.

However, allowing $\omega_s \neq 0$ in (52) enables the underlying trend in $\widehat{y}_{T+h|T}$ to be constant over seasons. Specifically, $\tau_s = \tau$ $(s = 1, \ldots, S)$ requires $\omega_s = (1 - \alpha_s)\tau$, which implies an intercept in (52) whose value is restricted over $s = 1, \ldots, S$. The interpretation is that the trend in the periodic difference $(1 - \alpha_s L)\widehat{y}_{T+h|T}$ must counteract the diverging trends that would otherwise arise in the forecasts $\widehat{y}_{T+h|T}$ over seasons; see Paap and Franses (1999) or Ghysels and Osborn (2001, pp. 155–156). An important implication is that if forecasts with diverging trends over seasons are implausible, then a constant (nonzero) trend can be achieved through the imposition of appropriate restrictions on the trend terms in the forecast function for the PIAR model.

[5] Stationarity for the periodic component here requires only $|\beta_1 \beta_2 \cdots \beta_S| < 1$.

3.4. Forecasting with misspecified models

Despite their theoretical attractions in some economic contexts, periodic models are not widely used for forecasting in economics. Therefore, it is relevant to consider the implications of applying an ARMA forecasting model to periodic GDP. This question is studied by Osborn (1991), building on Tiao and Grupe (1980).

It is clear from (44) and (45) that the autocovariances of a stationary PAR process differ over seasons. Denoting the autocovariance for season s at lag k by $\gamma_{sk} = E(x_{Sn+s}x_{Sn+s-k})$, the overall mean autocovariance at lag k is

$$\gamma_k = \frac{1}{S}\sum_{s=1}^{S} \gamma_{sk}. \tag{53}$$

When an ARMA model is fitted, asymptotically it must account for all nonzero autocovariances γ_k, $k = 0, 1, 2, \ldots$. Using (53), Tiao and Grupe (1980) and Osborn (1991) show that the implied ARMA model fitted to a PAR(p) process has, in general, a purely seasonal autoregressive operator of order p, together with a potentially high order moving average.

As a simple case, consider a purely stochastic PAR(1) process for $S = 2$ seasons per year, so that

$$\begin{aligned} x_{Sn+s} &= \phi_s x_{Sn+s-1} + \varepsilon_{Sn+s} \\ &= \phi_1\phi_2 x_{Sn+s-2} + \varepsilon_{Sn+s} + \phi_{s-1}\varepsilon_{Sn+s-1}, \quad s = 1, 2 \end{aligned} \tag{54}$$

where white noise ε_{Sn+s} has $E(\varepsilon^2_{Sn+s}) = \sigma^2_s$ and $\phi_0 = \phi_2$. The corresponding misspecified ARMA model that accounts for the autocovariances (53) effectively takes a form of average across the two processes in (54) to yield

$$x_{Sn+s} = \phi_1\phi_2 x_{Sn+s-2} + u_{Sn+s} + \theta u_{Sn+s-1} \tag{55}$$

where u_{Sn+s} has autocovariances $\gamma_k = 0$ for all lags $k = 1, 2 \ldots$. From known results concerning the accuracy of forecasting using aggregate and disaggregate series, the MSFE at any horizon h using the (aggregate) ARMA representation (54) must be at least as large as the mean MSFE over seasons for the true (disaggregate) PAR(1) process.

As in the analysis of misspecified processes in the discussion of linear models in the previous section, these results take no account of estimation effects. To the extent that, in practice, periodic models require the estimation of more coefficients than ARMA ones, the theoretical forecasting advantage of the former over the latter for a true periodic DGP will not necessarily carry over when observed data are employed.

3.5. Periodic cointegration

Periodic cointegration relates to cointegration between individual processes that are either periodically integrated or seasonally integrated. To concentrate on the essential issues, we consider periodic cointegration between the univariate nonstationary process y_{Sn+s} and the vector nonstationary process x_{Sn+s} as implying that

$$z_{Sn+s} = y_{Sn+s} - \alpha_s' x_{Sn+s}, \quad s = 1, \dots, S, \tag{56}$$

is a (possibly periodic) stationary process, with not all vectors α_s equal over $s = 1, \dots, S$. The additional complications of so-called partial periodic cointegration will not be considered. We also note that there has been much confusion in the literature on periodic processes relating to types of cointegration that can apply. These issues are discussed by Ghysels and Osborn (2001, pp. 168–171).

In both theoretical developments and empirical applications, the most popular single equation periodic cointegration model [PCM] has the form:

$$\begin{aligned}\Delta_S y_{Sn+s} = {} & \sum_{s=1}^{S} \mu_s D_{s,Sn+s} + \sum_{s=1}^{S} \lambda_s D_{s,Sn+s} \left(y_{Sn+s-S} - \alpha_s' x_{Sn+s-S}\right) \\ & + \sum_{k=1}^{p} \phi_k \Delta_S y_{Sn+s-k} + \sum_{k=0}^{p} \delta_k' \Delta_S x_{Sn+s-k} + \varepsilon_{Sn+s}\end{aligned} \tag{57}$$

where y_{Sn+s} is the variable of specific interest, x_{Sn+s} is a vector of weakly exogenous explanatory variables and ε_{Sn+s} is white noise. Here λ_s and α_s' are seasonally varying adjustment and long-run parameters, respectively; the specification of (57) could allow the disturbance variance to vary over seasons. As discussed by Ghysels and Osborn (2001, p. 171) this specification implicitly assumes that the individual variables of y_{Sn+s}, x_{Sn+s} are seasonally integrated, rather than periodically integrated.

Boswijk and Franses (1995) develop a Wald test for periodic cointegration through the unrestricted model

$$\begin{aligned}\Delta_S y_{Sn+s} = {} & \sum_{s=1}^{S} \mu_s D_{s,Sn+s} + \sum_{s=1}^{S} \left(\delta_{1s} D_{s,Sn+s} y_{Sn+s-S} + \delta_{2s}' D_{s,Sn+s} x_{Sn+s-4}\right) \\ & + \sum_{k=1}^{p} \beta_k \Delta_S y_{Sn+s-k} + \sum_{k=0}^{p} \tau_k' \Delta_S x_{Sn+s-k} + \varepsilon_{Sn+s}\end{aligned} \tag{58}$$

where under cointegration $\delta_{1s} = \lambda_s$ and $\delta_{2s} = -\alpha_s' \lambda_s$. Defining $\delta_s = (\delta_{1s}, \delta_{2s}')'$ and $\delta = (\delta_1', \delta_2', \dots, \delta_S')'$, the null hypothesis of no cointegration in any season is given by H_0: $\delta = 0$. Because cointegration for one season s does not necessarily imply cointegration for all $s = 1, \dots, S$, the alternative hypothesis H_1: $\delta \neq 0$ implies cointegration for at least one s. Relevant critical values for the quarterly case are given in Boswijk and Franses (1995), who also consider testing whether cointegration applies in individual seasons and whether cointegration is nonperiodic.

Since periodic cointegration is typically applied in contexts that implicitly assume seasonally integrated variables, it seems obvious that the possibility of seasonal cointegration should also be considered. Although Franses (1993, 1995) and Ghysels and

Osborn (2001, pp. 174–176) make some progress towards a testing strategy to distinguish between periodic and seasonal cointegration, this issue has yet to be fully worked out in the literature.

When the periodic ECM model of (57) is used for forecasting, a separate model is (of course) required to forecast the weakly exogenous variables in x.

3.6. *Empirical forecast comparisons*

Empirical studies of the forecast performance of periodic models for economic variables are mixed. Osborn and Smith (1989) find that periodic models produce more accurate forecasts than nonperiodic ones for the major components of quarterly UK consumers expenditure. However, although Wells (1997) finds evidence of periodic coefficient variation in a number of US time series, these models do not consistently produce improved forecast accuracy compared with nonperiodic specifications. In investigating the forecasting performance of PAR models, Rodrigues and Gouveia (2004) observe that using parsimonious periodic autoregressive models, with fewer separate "seasons" modelled than indicated by the periodicity of the data, presents a clear advantage in forecasting performance over other models. When examining forecast performance for observed UK macroeconomic time series, Novales and Flores de Fruto (1997) draw a similar conclusion.

As noted in our previous discussion, the role of deterministic variables is important in periodic models. Using the same series as Osborn and Smith (1989), Franses and Paap (2002) consider taking explicit account of the appropriate form of deterministic variables in PAR models and adopt encompassing tests to formally evaluate forecast performance.

Relatively few studies consider the forecast performance of periodic cointegration models. However, Herwartz (1997) finds little evidence that such models improve accuracy for forecasting consumption in various countries, compared with constant parameter specifications. In comparing various vector systems, Löf and Franses (2001) conclude that models based on seasonal differences generally produce more accurate forecasts than those based on first differences or periodic specifications.

In view of their generally unimpressive performance in empirical forecast comparisons to date, it seems plausible that parsimonious approaches to periodic ECM modelling may be required for forecasting, since an unrestricted version of (57) may imply a large number of parameters to be estimated. Further, as noted in the previous section, there has been some confusion in the literature about the situations in which periodic cointegration can apply and there is no clear testing strategy to distinguish between seasonal and periodic cointegration. Clarification of these issues may help to indicate the circumstances in which periodic specifications yield improved forecast accuracy over nonperiodic models.

4. Other specifications

The previous sections have examined linear models and periodic models, where the latter can be viewed as linear models with a structure that changes with the season. The simplest models to specify and estimate are linear (time-invariant) ones. However, there is no *a priori* reason why seasonal structures should be linear and time-invariant. The preferences of economic agents may change over time or institutional changes may occur that cause the seasonal pattern in economic variables to alter in a systematic way over time or in relation to underlying economic conditions, such as the business cycle.

In recent years a burgeoning literature has examined the role of nonlinear models for economic modelling. Although much of this literature takes the context as being nonseasonal, a few studies have also examined these issues for seasonal time series. Nevertheless, an understanding of the nature of change over time is a fundamental prerequisite for accurate forecasting.

The present section first considers nonlinear threshold and Markov switching time series models, before turning to a notion of seasonality different from that discussed in previous sections, namely seasonality in variance. Consider for expository purposes the general model,

$$y_{Sn+s} = \mu_{Sn+s} + \xi_{Sn+s} + x_{Sn+s}, \tag{59}$$

$$\psi(L)x_{Sn+s} = \varepsilon_{Sn+s} \tag{60}$$

where μ_{Sn+s} and ξ_{Sn+s} represent deterministic variables which will be presented in detail in the following sections, $\varepsilon_{Sn+s} \sim \Gamma(0, h_t)$, Γ is a probability distribution and h_t represents the assumed variance which can be constant over time or time varying.

In the following section we start to look at nonlinear models and the implications of seasonality in the mean, which will be introduced through μ_{Sn+s} and ξ_{Sn+s}, considering that the errors are i.i.d. $N(0, \sigma^2)$; and in Section 4.2 proceed to investigate the modelling of seasonality in variance, considering that the errors follow GARCH or stochastic volatility type behaviour and allowing for the seasonal behavior in volatility to be deterministic and stochastic.

4.1. Nonlinear models

Although many different types of nonlinear models have been proposed, perhaps those used in a seasonal context are of the threshold or regime-switching types. In both cases, the relationship is assumed to be linear within a regime. These nonlinear models focus on the interaction between seasonality and the business cycle, since Ghysels (1994b), Canova and Ghysels (1994), Matas-Mir and Osborn (2004) and others have shown that these are interrelated.

4.1.1. Threshold seasonal models

In this class of models, the regimes are defined by the values of some variable in relation to specific thresholds, with the transition between regimes being either abrupt or smooth. To distinguish these, the former are referred to as threshold autoregressive (TAR) models, while the latter are known as smooth transition autoregressive (STAR) models. Threshold models have been applied to seasonal growth in output, with the annual output growth used as the business cycle indicator.

Cecchitti and Kashyap (1996) provide some theoretical basis for an interaction between seasonality and the business cycle, by outlining an economic model of seasonality in production over the business cycle. Since firms may hit capacity restrictions when production is high, they will reallocate production to the usually slack summer months near business cycle peaks.

Motivated by this hypothesis, Matas-Mir and Osborn (2004) consider the seasonal TAR model for monthly data given as

$$\begin{aligned}\Delta_1 y_{Sn+s} = {} & \mu_0 + \eta_0 I_{Sn+s} + \tau_0(Sn+s) \\ & + \sum_{j=1}^{S}\bigl[\mu_j^* + \eta_j^* I_{Sn+s} + \tau_j^*(Sn+s)\bigr]D_{j,Sn+s}^* \\ & + \sum_{i=1}^{p}\phi_i \Delta_1 y_{Sn+s-i} + \varepsilon_{Sn+s}\end{aligned} \tag{61}$$

where $S = 12$, $\varepsilon_{Sn+s} \sim \text{iid}(0,\sigma^2)$, $D_{j,Sn+s}^*$ is a seasonal dummy variable and the regime indicator I_{Sn+s} is defined in terms of a threshold value r for the lagged annual change in y. Note that this model results from (59) and (60) by considering that $\mu_{Sn+s} = \delta_0+\gamma_0(Sn+s)+\sum_{j=1}^{S}[\delta_j+\gamma_j(Sn+s)]D_{j,Sn+s}$, $\xi_{Sn+s} = [\alpha_0+\sum_{j=1}^{S}\alpha_j D_{j,Sn+s}]I_{Sn+s}$ and $\psi(L) = \phi(L)\Delta_1$ is a polynomial of order $p+1$. The nonlinear specification of (61) allows the overall intercept and the deterministic seasonality to change with the regime, but (for reasons of parsimony) not the dynamics. Systematic changes in seasonality are permitted through the inclusion of seasonal trends. Matas-Mir and Osborn (2004) find support for the seasonal nonlinearities in (61) for around 30 percent of the industrial production series they analyze for OECD countries.

A related STAR specification is employed by van Dijk, Strikholm and Terasvirta (2003). However, rather than using a threshold specification which results from the use of the indicator function I_{Sn+s}, these authors specify the transition between regimes using the logistic function

$$G_i(\varphi_{it}) = \bigl[1+\exp\{-\gamma_i(\varphi_{it}-c_i)/\sigma_{s_{it}}\}\bigr]^{-1}, \quad \gamma_i > 0 \tag{62}$$

for a transition variable φ_{it}. In fact, they allow two such transition functions ($i = 1, 2$) when modelling the quarterly change in industrial production for G7 countries, with one transition variable being the lagged annual change ($\varphi_{1t} = \Delta_4 y_{t-d}$ for some delay d), which can be associated with the business cycle, and the other transition variable being time ($\varphi_{2t} = t$). Potentially all coefficients, relating to both the seasonal dummy variables and the autoregressive dynamics are allowed to change with the regime. These authors conclude that changes in the seasonal pattern associated with the time transition are more important than those associated with the business cycle.

In a nonseasonal context, Clements and Smith (1999) investigate the multi-step forecast performance of TAR models via empirical MSFEs and show that these models perform significantly better than linear models particularly in cases when the forecast origin covers a recession period. It is notable that recessions have fewer observations than expansions, so that their forecasting advantage appears to be in atypical periods.

There has been little empirical investigation of the forecast accuracy of nonlinear seasonal threshold models for observed series. The principal available study is Franses and van Dijk (2005), who consider various models of seasonality and nonlinearity for quarterly industrial production for 18 OECD countries. They find that, in general, linear models perform best at short horizons, while nonlinear models with more elaborate seasonal specifications are preferred at longer horizons.

4.1.2. Periodic Markov switching regime models

Another approach to model the potential interaction between seasonal and business cycles is through periodic Markov switching regime models. Special cases of this class include the (aperiodic) switching regime models considered by Hamilton (1989, 1990), among many others. Ghysels (1991, 1994b, 1997) presented a periodic Markov switching structure which was used to investigate the nonuniformity over months of the distribution of the NBER business cycle turning points for the US. The discussion here, which is based on Ghysels (2000) and Ghysels, Bac and Chevet (2003), will focus first on a simplified illustrative example to present some of the key features and elements of interest. The main purpose is to provide intuition for the basic insights. In particular, one can map periodic Markov switching regime models into their linear representations. Through the linear representation one is able to show that hidden periodicities are left unexploited and can potentially improve forecast performance.

Consider a univariate time series process, again denoted $\{y_{Sn+s}\}$. It will typically represent a growth rate of, say, GNP. Moreover, for the moment, it will be assumed the series does not exhibit seasonality in the mean (possibly because it was seasonally adjusted) and let $\{y_{Sn+s}\}$ be generated by the following stochastic structure:

$$\bigl(y_{Sn+s} - \mu\bigl[(i_{Sn+s}, \mathbf{v})\bigr]\bigr) = \phi\bigl(y_{Sn+s-1} - \mu\bigl[(i_{Sn+s-1}, \mathbf{v} - \mathbf{1})\bigr]\bigr) + \varepsilon_{Sn+s} \tag{63}$$

where $|\phi| < 1$, ε_t is i.i.d. $N(0, \sigma^2)$ and $\mu[\cdot]$ represents an intercept shift function. If $\mu \equiv \bar{\mu}$, i.e., a constant, then (63) is a standard linear stationary Gaussian AR(1) model. Instead, following Hamilton (1989), we assume that the intercept changes according to a Markovian switching regime model. However, in (63) we have $x_t \equiv (i_t, \mathbf{v})$, namely, the state of the world is described by a stochastic switching regime process $\{i_t\}$ and a seasonal indicator process $\mathbf{v}$. The $\{i_{Sn+s}\}$ and $\{\mathbf{v}\}$ processes interact in the following way, such that for $i_{Sn+s} \in \{0, 1\}$:[6]

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & q(\mathbf{v}) & 1-q(\mathbf{v}) \\ 1 & 1-p(\mathbf{v}) & p(\mathbf{v}) \end{array} \tag{64}$$

where the transition probabilities $q(\cdot)$ and $p(\cdot)$ are allowed to change with the season. When $p(\cdot) = \bar{p}$ and $q(\cdot) = \bar{q}$, we obtain the standard homogeneous Markov chain model considered by Hamilton. However, if for at least one season the transition probability matrix differs, we have a situation where a regime shift will be more or less likely depending on the time of the year. Since $i_{Sn+s} \in \{0, 1\}$, consider the mean shift function:

$$\mu[(i_t, \mathbf{v})] = \alpha_0 + \alpha_1 i_{Sn+s}, \quad \alpha_1 > 0. \tag{65}$$

Hence, the process $\{y_{Sn+s}\}$ has a mean shift α_0 in state 1 ($i_{Sn+s} = 0$) and $\alpha_0 + \alpha_1$ in state 2. These above equations are a version of Hamilton's model with a periodic stochastic switching process. If state 1 with low mean drift is called a recession and state 2 an expansion, then we stay in a recession or move to an expansion with a probability scheme that depends on the season.

The structure presented so far is relatively simple, yet as we shall see, some interesting dynamics and subtle interdependencies emerge. It is worth comparing the AR(1) model with a periodic Markovian stochastic switching regime structure and the more conventional linear ARMA processes as well as periodic ARMA models. Let us perhaps start by briefly explaining intuitively what drives the connections between the different models. The model with y_{Sn+s} typically representing a growth series, is covariance stationary under suitable regularity conditions discussed in Ghysels (2000). Consequently, the process has a linear Wold MA representation. Yet, the time series model provides a relatively parsimonious structure which determines nonlinearly predictable MA innovations. In fact, there are two layers beneath the Wold MA representation. One layer relates to *hidden periodicities*, as described in Tiao and Grupe (1980) or Hansen and Sargent (1993), for instance. Typically, such hidden periodicities can be uncovered via augmentation of the state space with the augmented system having a linear representation. However, the periodic switching regime model imposes *further structure* even after the hidden periodicities are uncovered. Indeed, there is a second layer which makes the innovations of the augmented system nonlinearly predictable. Hence, the model has nonlinearly predictable innovations and features of hidden periodicities combined.

To develop this more explicitly, let us first note that the switching regime process $\{i_{Sn+s}\}$ admits the following AR(1) representation:

$$i_{Sn+s} = \left[1 - q(\mathbf{v}_t)\right] + \lambda(\mathbf{v}_t) i_{t-1} + v_{Sn+s}(\mathbf{v}) \tag{66}$$

[6] In order to avoid too cumbersome notation, we did not introduce a separate notation for the theoretical representation of stochastic processes and their actual realizations.

where $\lambda(\cdot) \in \{\lambda^1, \ldots, \lambda^{\mathcal{S}}\}$ with $\lambda(\mathbf{v}) \equiv -1 + p(\mathbf{v}) + q(\mathbf{v}) = \lambda^{\mathrm{s}}$ for $\mathbf{v} = \mathbf{v}$. Moreover, conditional on $i_{t-1} = 1$,

$$v_{Sn+s}(\mathbf{v}) = \begin{cases} \big(1 - p(\mathbf{v})\big) & \text{with probability } p(\mathbf{v}), \\ -p(\mathbf{v}) & \text{with probability } 1 - p(\mathbf{v}_t) \end{cases} \tag{67}$$

while conditional on $i_{t-1} = 0$,

$$v_{Sn+s}(\mathbf{v}) = \begin{cases} -\big(1 - q(\mathbf{v})\big) & \text{with probability } q(\mathbf{v}), \\ q(\mathbf{v}) & \text{with probability } 1 - q(\mathbf{v}_t). \end{cases} \tag{68}$$

Equation (66) is a periodic AR(1) model where all the parameters, including those governing the error process, may take on different values every season. Of course, this is a different way of saying that the "state-of-the-world" is not only described by $\{i_{Sn+s}\}$ but also $\{\mathbf{v}\}$. While (66) resembles the periodic ARMA models which were discussed by Tiao and Grupe (1980), Osborn (1991) and Hansen and Sargent (1993), among others, it is also fundamentally different in many respects. The most obvious difference is that the innovation process has a discrete distribution.

The linear time invariant representation for the stochastic switching regime process i_{Sn+s} is a finite order ARMA process, as we shall explain shortly. One should note that the process will certainly not be represented by an AR(1) process as it will not be Markovian in such a straightforward way when it is expressed by a univariate AR(1) process, since part of the state space is "missing". A more formal argument can be derived directly from the analysis in Tiao and Grupe (1980) and Osborn (1991).[7] The periodic nature of autoregressive coefficients pushes the seasonality into annual lags of the AR polynomial and substantially complicates the MA component.

Ultimately, we are interested in the time series properties of $\{y_{Sn+s}\}$. Since

$$y_{Sn+s} = \alpha_0 + \alpha_1 i_{Sn+s} + (1 - \phi L)^{-1} \varepsilon_{Sn+s}, \tag{69}$$

and ε_{Sn+s} was assumed Gaussian and independent, we can simply view $\{y_{Sn+s}\}$ as the sum of two independent unobserved processes: namely, $\{i_{Sn+s}\}$ and the process $(1 - \phi L)^{-1}\varepsilon_{Sn+s}$. Clearly, all the features just described about the $\{i_{Sn+s}\}$ process will be translated into similar features inherited by the observed process y_{Sn+s}, while y_{Sn+s} has the following linear time series representation:

$$w_y(z) = \alpha_1^2 w_i(z) + 1/\big[(1 - \phi z)\big(1 - \phi z^{-1}\big)\big]\sigma^2/2\pi. \tag{70}$$

This linear representation has hidden periodic properties and a stacked skip sampled version of the $(1 - \phi L)^{-1}\varepsilon_{Sn+s}$ process. Finally, the vector representation obtained as such would inherit the nonlinear predictable features of $\{i_{Sn+s}\}$.

[7] Osborn (1991) establishes a link between periodic processes and contemporaneous aggregation and uses it to show that the periodic process must have an average forecast MSE at least as small as that of its univariate time invariant counterpart. A similar result for periodic hazard models and scoring rules for predictions is discussed in Ghysels (1993).

Let us briefly return to (69). We observe that the linear representation has seasonal mean shifts that appear as a "deterministic seasonal" in the univariate representation of y_{Sn+s}. Hence, besides the spectral density properties in (70), which may or may not show peaks at the seasonal frequency, we note that periodic Markov switching produces seasonal mean shifts in the univariate representation. This result is, of course, quite interesting since intrinsically we have a purely random stochastic process with occasional mean shifts. The fact that we obtain something that resembles a deterministic seasonal simply comes from the unequal propensity to switch regime (and hence mean) during some seasons of the year.

4.2. Seasonality in variance

So far our analysis has concentrated on models which account for seasonality in the conditional mean only, however a different concept of considerable interest, particularly in the finance literature, is the notion of seasonality in the variance. There is both seasonal heteroskedasticity in daily data and intra-daily data. For daily data, see for instance Tsiakas (2004b). For intra-daily see, e.g., Andersen and Bollerslev (1997). In a recent paper, Martens, Chang and Taylor (2002) present evidence which shows that explicitly modelling intraday seasonality improves out-of-sample forecasting performance; see also Andersen, Bollerslev and Lange (1999).

The notation needs to be slightly generalized in order to handle intra-daily seasonality. In principle we could have three subscripts, like for instance m, s, and n, referring to the mth intra-day observation in 'season' s (e.g., week s) in year n. Most often we will only use m and T, the latter being the total sample. Moreover, since seasonality is often based on daily observations we will often use d as a subscript to refer to a particular day (with m intra-daily observations).

In order to investigate whether out-of-sample forecasting is improved when using seasonal methods, Martens, Chang and Taylor (2002) consider a conventional t-distribution GARCH(1,1) model as benchmark

$$\begin{aligned} r_t &= \mu + \varepsilon_t, \\ \varepsilon_t | \Psi_{t-1} &\sim D(0, h_t), \\ h_t &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \end{aligned}$$

where Ψ_{t-1} corresponds to the information set available at time $t-1$ and D represents a scaled t-distribution. In this context, the out-of-sample variance forecast is given by

$$\widehat{h}_{T+1} = \widehat{\omega} + \widehat{\alpha} \varepsilon_T^2 + \widehat{\beta} h_T. \tag{71}$$

As Martens, Chang and Taylor (2002) also indicate, for GARCH models with conditional scaled *t-distributions* with υ degrees of freedom, the expected absolute return is given by

$$E|r_{T+1}| = 2 \frac{\sqrt{\upsilon - 2}}{\sqrt{\pi}} \frac{\Gamma[(\upsilon + 1)/2]}{\Gamma[\upsilon/2](\upsilon - 1)} \sqrt{\widehat{h}_{T+1}}$$

where Γ is the gamma-function.

However, as pointed out by Andersen and Bollerslev (1997, p. 125), standard ARCH modelling implies a geometric decay in the autocorrelation structure and cannot accommodate strong regular cyclical patterns. In order to overcome this problem, Andersen and Bollerslev suggest a simple specification of interaction between the pronounced intraday periodicity and the strong daily conditional heteroskedasticity as

$$r_t = \sum_{m=1}^{M} r_{t,m} = \sigma_t \frac{1}{M^{1/2}} \sum_{m=1}^{M} v_m Z_{t,m} \tag{72}$$

where r_t denotes the daily continuous compounded return calculated from the M uncorrelated intraday components $r_{t,m}$, σ_t denotes the conditional volatility factor for day t, v_m represents the deterministic intraday pattern and $Z_{t,m} \sim \text{iid}(0, 1)$, which is assumed to be independent of the daily volatility process $\{\sigma_t\}$. Both volatility components must be non-negative, i.e., $\sigma_t > 0$ a.s. for all t and $v_m > 0$ for all m.

4.2.1. Simple estimators of seasonal variances

In order to take into account the intradaily seasonal pattern, Taylor and Xu (1997) consider for each intraday period the average of the squared returns over all trading days, i.e., the variance estimate is given as

$$v_m^2 = \frac{1}{D} \sum_{t=1}^{N} r_{t,m}^2, \quad n = 1, \ldots, M \tag{73}$$

where N is the number of days. An alternative is to use

$$v_{d,m}^2 = \frac{1}{M_d} \sum_{k \in T_d} r_{k,m}^2$$

where T_d is the set of daily time indexes that share the same day of the week as time index d, and M_d is the number of time indexes in T_d. Note that this approach, in contrast to (73), takes into account the day of the week. Following the assumption that volatility is the product of seasonal volatility and a time-varying nonseasonal component as in (72), the seasonal variances can be computed as

$$v_{d,m}^2 = \exp\left[\frac{1}{M_d} \sum_{k \in T_d} \ln\bigl((r_{k,m} - \bar{r})^2\bigr)\right]$$

where $\bar{r}$ is the overall mean taken over all returns.

The purpose of estimating these seasonal variances is to scale the returns,

$$\widetilde{r}_t \equiv \widetilde{r}_{d,m} \equiv \frac{r_{d,m}}{v_{d,m}}$$

in order to estimate a conventional GARCH(1,1) model for the scaled returns, and hence, forecasts of $\widetilde{h}_{T+1}$ can be obtained in the conventional way as in (71). To transform the volatility forecasts for the scaled returns into volatility forecasts for the original returns, Martens, Chang and Taylor (2002) suggest multiplying the volatility forecasts by the appropriate estimate of the seasonal standard deviation, $v_{d,m}$.

4.2.2. *Flexible Fourier form*

The Flexible Fourier Form (FFF) [see Gallant (1981)] is a different approach to capture deterministic intraday volatility pattern; see *inter alia* Andersen and Bollerslev (1997, p. 152) and Beltratti and Morana (1999). Andersen and Bollerslev assume that the intraday returns are given as

$$r_{d,m} = E(r_{d,m}) + \frac{\sigma_d v_{d,m} Z_{d,m}}{M^{1/2}} \tag{74}$$

where $E(r_{d,m})$ denotes the unconditional mean and $Z_{d,m} \sim \text{iid}(0, 1)$. From (74) they define the variable

$$x_{d,m} \equiv 2\ln\bigl[\bigl|r_{d,m} - E(r_{d,m})\bigr|\bigr] - \ln\sigma_d^2 + \ln M = \ln v_{d,m}^2 + \ln Z_{d,m}^2.$$

Replacing $E(r_{d,m})$ by the sample average of all intraday returns and σ_d by an estimate from a daily volatility model, $\widehat{x}_{d,m}$ is obtained. Treating $\widehat{x}_{d,m}$ as dependent variable, the seasonal pattern is obtained by OLS as

$$\begin{aligned}\widehat{\widehat{x}}_{d,m} \equiv \sum_{j=0}^{J} \sigma_d^j \Biggl[\mu_{0j} + \mu_{1j}\frac{m}{M_1} + \mu_{2j}\frac{n^2}{M_2} + \sum_{i=1}^{l} \lambda_{ij} I_{t=d_t} \\ + \sum_{i=1}^{p}\left(\gamma_{ij}\cos\frac{2\pi i n}{M} + \delta_{ij}\sin\frac{2\pi i n}{M}\right)\Biggr],\end{aligned}$$

where $M_1 = (M+1)/2$ and $M_2 = (M+1)(M+2)/6$ are normalizing constants and p is set equal to four. Each of the corresponding $J+1$ FFFs are parameterized by a quadratic component (the terms with μ coefficients) and a number of sinusoids. Moreover, it may be advantageous to include time-specific dummies for applications in which some intraday intervals do not fit well within the overall regular periodic pattern (the λ coefficients).

Hence, once $\widehat{\widehat{x}}_{d,m}$ is estimated, the intraday seasonal volatility pattern can be determined as [see Martens, Chang and Taylor (2002)],

$$\widehat{v}_{d,m} = \exp\bigl(\widehat{\widehat{x}}_{d,m}/2\bigr)$$

or alternatively [as suggested by Andersen and Bollerslev (1997, p. 153)],

$$\widehat{v}_{d,m} = \frac{T\exp(\widehat{\widehat{x}}_{d,m}/2)}{\sum_{d=1}^{[T/M]}\sum_{n=1}^{M}\exp(\widehat{\widehat{x}}_{d,m}/2)}$$

which results from the normalization $\sum_{d=1}^{[T/M]} \sum_{n=1}^{M} v_{d,m} \equiv 1$, where $[T/M]$ represents the number of trading days in the sample.

4.2.3. Stochastic seasonal pattern

The previous two subsections assume that the observed seasonal pattern is deterministic. However, there may be no reason that justifies daily or weekly seasonal behavior in volatility as deterministic. Beltratti and Morana (1999) provide, among other things, a comparison between deterministic and stochastic models for the filtering of high frequency returns. In particular, the deterministic seasonal model of Andersen and Bollerslev (1997), described in the previous subsection, is compared with a model resulting from the application of the structural methodology developed by Harvey (1994).

The model proposed by Beltratti and Morana (1999) is an extension of one introduced by Harvey, Ruiz and Shephard (1994), who apply a stochastic volatility model based on the structural time series approach to analyze daily exchange rate returns. This methodology is extended by Payne (1996) to incorporate an intra-day fixed seasonal component, whereas Beltratti and Morana (1999) extend it further to accommodate stochastic intra-daily cyclical components, as

$$r_{t,m} = \bar{r}_{t,m} + \sigma_{t,m}\varepsilon_{t,m} = \bar{r}_{t,m} + \sigma\varepsilon_{t,m} \exp\left(\frac{\mu_{t,m} + h_{t,m} + c_{t,m}}{2}\right) \tag{75}$$

for $t = 1, \ldots, T$, $n = 1, \ldots, M$; and where σ is a scale factor, $\varepsilon_{t,m} \sim \text{iid}(0, 1)$, $\mu_{t,m}$ is the non-stationary volatility component given as $\mu_{t,m} = \mu_{t,m-1} + \xi_{t,m}$, $\xi_{t,m} \sim \text{nid}(0, \sigma_\xi^2)$, $h_{t,m}$ is the stochastic stationary acyclic volatility component, $h_{t,m} = \phi h_{t,m-1} + \vartheta_{t,m}$, $\vartheta_{t,m} \sim \text{nid}(0, \sigma_\vartheta^2)$, $|\phi| < 1$, c_t is the cyclical volatility component and $\bar{r}_{t,m} = E[r_{t,m}]$.

As suggested by Beltratti and Morana, squaring both sides and taking logs, allows (75) to be rewritten as,

$$\ln\bigl(|r_{t,m} - \bar{r}_{t,m}|\bigr)^2 = \ln\left[\sigma\varepsilon_{t,m} \exp\left(\frac{\mu_{t,m} + h_{t,m} + c_{t,m}}{2}\right)\right]^2,$$

that is,

$$2\ln|r_{t,m} - \bar{r}_{t,m}| = \iota + \mu_{t,m} + h_{t,m} + c_{t,m} + w_{t,m}$$

where $\iota = \ln\sigma^2 + E[\ln\varepsilon_{t,m}^2]$ and $w_{t,m} = \ln\varepsilon_{t,m}^2 - E[\ln\varepsilon_{t,m}^2]$.

The c_t component is broken into a number of cycles corresponding to the fundamental daily frequency and its intra-daily harmonics, i.e., $c_{t,m} = \sum_{i=1}^{2} c_{i,t,m}$. Beltratti and Morana model the fundamental daily frequency, $c_{1,t,m}$, as stochastic while its harmonics, $c_{2,t,m}$, as deterministic. In other words, following Harvey (1994), the stochastic cyclical component, $c_{1,t,m}$, is considered in state space form as

$$c_{1,t,m} = \begin{bmatrix} \psi_{1,t,m} \\ \psi_{1,t,m}^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \psi_{1,t,m-1} \\ \psi_{1,t,m-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_{1,t,m} \\ \kappa_{1,t,m}^* \end{bmatrix}$$

where $0 \leqslant \rho \leqslant 1$ is a damping factor and $\kappa_{1,t,m} \sim \text{nid}(0, \sigma^2_{1,\kappa})$ and $\kappa^*_{1,t,m} \sim \text{nid}(0, \sigma^{*2}_{1,\kappa})$ are white noise disturbances with $Cov(\kappa_{1,t,m}, \kappa^*_{1,t,m}) = 0$. Whereas, $c_{2,t,m}$ is modelled using a flexible Fourier form as

$$c_{2,t,m} = \mu_1 \frac{m}{M_1} + \mu_2 \frac{n^2}{M_2} + \sum_{i=2}^{p} (\delta_{ci} \cos i\lambda n + \delta_{si} \sin i\lambda n).$$

It can be observed from the specification of these components that this model encompasses that of Andersen and Bollerslev (1997).

One advantage of this state space formulation results from the possibility that the various components may be estimated simultaneously. One important conclusion that comes out of the empirical evaluation of this model, is that it presents some superior results when compared with the models that treat seasonality as strictly deterministic; for more details see Beltratti and Morana (1999).

4.2.4. Periodic GARCH models

In the previous section we dealt with intra-daily returns data. Here we return to daily returns and to daily measures of volatility. An approach to seasonality considered by Bollerslev and Ghysels (1996) is the periodic GARCH (P-GARCH) model which is explicitly designed to capture (daily) seasonal time variation in the second-order moments; see also Ghysels and Osborn (2001, pp. 194–198). The P-GARCH includes all GARCH models in which hourly dummies, for example, are used in the variance equation.

Extending the information set Ψ_{t-1} with a process defining the stage of the periodic cycle at each point, say to Ψ^s_{t-1}, the P-GARCH model is defined as,

$$\begin{aligned} r_t &= \mu + \varepsilon_t, \\ \varepsilon_t | \Psi^s_{t-1} &\sim D(0, h_t), \\ h_t &= \omega_{s(t)} + \alpha_{s(t)} \varepsilon^2_{t-1} + \beta_{s(t)} h_{t-1} \end{aligned} \tag{76}$$

where $s(t)$ refers to the stage of the periodic cycle at time t. The periodic cycle of interest here is a repetitive cycle covering one week. Notice that there is resemblance with the periodic models discussed in Section 3.

The P-GARCH model is potentially more efficient than the methods described earlier. These methods [with the exception of Beltratti and Morana (1999)] first estimate the seasonals, and after deseasonalizing the returns, estimate the volatility of these adjusted returns. The P-GARCH model on the other hand, allows for simultaneous estimation of the seasonal effects and the remaining time-varying volatility.

As indicated by Ghysels and Osborn (2001, p. 195) in the existing ARCH literature, the modelling of non-trading day effects has typically been limited to $\omega_{s(t)}$, whereas (76) allows for a much richer dynamic structure. However, some caution is necessary as discussed in Section 3 for the PAR models, in order to avoid overparameterization.

Moreover, as suggested by Martens, Chang and Taylor (2002), one can consider the parameters $\omega_{s(t)}$ in (76) in such a way that they represent: (a) the average absolute/square returns (e.g., 240 dummies) or (b) the FFF. Martens, Chang and Taylor (2002) consider the second approach allowing for only one FFF for the entire week instead of separate FFF for each day of the week.

4.2.5. Periodic stochastic volatility models

Another popular class of models is the so-called stochastic volatility models [see, e.g., Ghysels, Harvey and Renault (1996) for further discussion]. In a recent paper Tsiakas (2004a) presents the periodic stochastic volatility (PSV) model. Models of stochastic volatility have been used extensively in the finance literature. Like GARCH-type models, stochastic volatility models are designed to capture the persistent and predictable component of daily volatility, however in contrast with GARCH models the assumption of a stochastic second moment introduces an additional source of risk.

The benchmark model considered by Tsiakas (2004a) is the conventional stochastic volatility model given as,

$$y_t = \alpha + \rho y_{t-1} + \eta_t \tag{77}$$

and

$$\eta_t = \varepsilon_t \upsilon_t, \quad \varepsilon_t \sim \text{nid}(0, 1)$$

where the persistence of the stochastic conditional volatility υ_t is captured by the latent log-variance process h_t, which is modelled as a dynamic Gaussian variable

$$\upsilon_t = \exp(h_t/2)$$

and

$$h_t = \mu + \beta' X_t + \phi(h_{t-1} - \mu) + \sigma \varrho_t, \quad \varrho_t \sim \text{nid}(0, 1). \tag{78}$$

Note that in this framework ε_t and ϱ_t are assumed to be independent and that returns and their volatility are stationary, i.e., $|\rho| < 1$ and $|\phi| < 1$, respectively.

Tsiakas (2004a) introduces a PSV model in which the constants (levels) in both the conditional mean and the conditional variances are generalized to account for day of the week, holiday (non-trading day) and month of the year effects.

5. Forecasting, seasonal adjustment and feedback

The greatest demand for forecasting seasonal time series is a direct consequence of removing seasonal components. The process, called seasonal adjustment, aims to filter raw data such that seasonal fluctuations disappear from the series. Various procedures exist and Ghysels and Osborn (2001, Chapter 4) provide details regarding the most commonly used, including the U.S. Census Bureau X-11 method and its recent upgrade, the X-12-ARIMA program and the TRAMO/SEATS procedure.

We cover three issues in this section. The first subsection discusses how forecasting seasonal time series is deeply embedded in the process of seasonal adjustment. The second handles forecasting of seasonally adjusted series and the final subsection deals with feedback and control.

5.1. Seasonal adjustment and forecasting

The foundation of seasonal adjustment procedures is the decomposition of a series into a trend cycle, and seasonal and irregular components. Typically a series y_t is decomposed into the *product* of a trend cycle y_t^{tc}, seasonal y_t^s, and irregular y_t^i. However, assuming the use of logarithms, we can consider the additive decomposition

$$y_t = y_t^{tc} + y_t^s + y_t^i. \tag{79}$$

Other decompositions exist [see Ghysels and Osborn (2001), Hylleberg (1986)], yet the above decomposition has been the focus of most of the academic research. Seasonal adjustment filters are two-sided, involving both leads and lags. The linear X-11 filter will serve the purpose here as illustrative example to explain the role of forecasting.[8] The linear approximation to the monthly X-11 filter is:

$$\begin{aligned}\nu_{X-11}^M(L) &= 1 - SM_C(L)M_2(L)\big\{1 - HM(L)\big[1 - SM_C(L)M_1(L)SM_C(L)\big]\big\}\\ &= 1 - SM_C(L)M_2(L) + SM_C(L)M_2(L)HM(L)\\ &\quad - SM_C^3(L)M_1(L)M_2(L)HM(L) + SM_C^3(L)M_1(L)M_2(L),\end{aligned} \tag{80}$$

where $SM_C(L) \equiv 1 - SM(L)$, a centered thirteen-term MA filter, namely $SM(L) \equiv (1/24)(1 + L)(1 + L + \cdots + L^{11})L^{-6}$, $M_1(L) \equiv (1/9)(L^S + 1 + L^{-S})^2$ with $S = 12$. A similar filter is the "3×5" seasonal moving average filter $M_2(L) \equiv (1/15)(\sum_{j=-1}^{1} L^{jS})(\sum_{j=-2}^{2} L^{jS})$ again with $S = 12$. The procedure also involves a $(2H + 1)$-term Henderson moving average filter $HM(L)$ [see Ghysels and Osborn (2001) the default value is $H = 6$, yielding a thirteen-term Henderson moving average filter].

The monthly X-11 filter has roughly 5 years of leads and lags. The original X-11 seasonal adjustment procedure consisted of an array of asymmetric filters that complemented the two-sided symmetric filter. There was a separate filter for each scenario of missing observations, starting with a concurrent adjustment filter when on past data and none of the future data. Each of the asymmetric filters, when compared to the symmetric filter, implicitly defined a forecasting model for the missing observations in the data. Unfortunately, these different asymmetric filters implied inconsistent forecasting

[8] The question whether seasonal adjustment procedures are, at least approximately, linear data transformations is investigated by Young (1968) and Ghysels, Granger and Siklos (1996).

models across time. To eliminate this inconsistency, a major improvement was designed and implemented by Statistics Canada and called X-11-ARIMA [Dagum (1980)] that had the ability to extend time series with forecasts and backcasts from ARIMA models prior to seasonal adjustment. As a result, the symmetric filter was always used and any missing observations were filled in with an ARIMA model-based prediction. Its main advantage was smaller revisions of seasonally adjusted series as future data became available [see, e.g., Bobbitt and Otto (1990)]. The U.S. Census Bureau also proceeded in 1998 to major improvements of the X-11 procedure. These changes were so important that they prompted the release of what is called X-12-ARIMA. Findley et al. (1998) provide a very detailed description of the new improved capabilities of the X-12-ARIMA procedure. It encompasses the improvements of Statistics Canada's X-11-ARIMA and encapsulates it with a front end regARIMA program, which handles regression and ARIMA models, and a set of diagnostics, which enhance the appraisal of the output from the original X-11-ARIMA. The regARIMA program has a set of built-in regressors for the monthly case [listed in Table 2 of Findley et al. (1998)]. They include a constant trend, deterministic seasonal effects, trading-day effects (for both stock and flow variables), length-of-month variables, leap year, Easter holiday, Labor day, and Thanksgiving dummy variables as well as additive outlier, level shift, and temporary ramp regressors.

Goméz and Maravall (1996) succeeded in building a seasonal adjustment package using signal extraction principles. The package consists of two programs, namely TRAMO (Time Series Regression with ARIMA Noise, Missing observations, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series). The TRAMO program fulfills the role of preadjustment, very much like regARIMA does for X-12-ARIMA adjustment. Hence, it performs adjustments for outliers, trading-day effects, and other types of intervention analysis [following Box and Tiao (1975)].

This brief description of the two major seasonal adjustment programs reveals an important fact: seasonal adjustment involves forecasting seasonal time series. The models that are used in practice are the univariate ARIMA models described in Section 2.

5.2. Forecasting and seasonal adjustment

Like it or not, many applied time series studies involve forecasting seasonally adjusted series. However, as noted in the previous subsection, pre-filtered data are predicted in the process of adjustment and this raises several issues. Further, due to the use of two-sided filters, seasonal adjustment of historical data involves the use of future values. Many economic theories rest on the behavioral assumption of rational expectations, or at least are very careful regarding the information set available to agents. In this regard the use of seasonally adjusted series may be problematic.

An issue rarely discussed in the literature is that forecasting seasonally adjusted series, should at least in principle be linked to the forecasting exercise that is embedded in the seasonal adjustment process. In the previous subsection we noted that since adjustment filters are two-sided, future realizations of the raw series have to be predicted.

Implicitly one therefore has a prediction model for the non-seasonal components y_t^{tc} and irregular y_t^i appearing in Equation (79). For example, how many unit roots is y_t^{tc} assumed to have when seasonal adjustment procedures are applied, and is the same assumption used when subsequently seasonally adjusted series are predicted? One might also think that the same time series model either implicitly or explicitly used for $y_t^{tc} + y_t^i$ should be subsequently used to predict the seasonally adjusted series. Unfortunately that is not the case, since the seasonally adjusted series equals $y_t^{tc} + y_t^i + e_t$, where the latter is an extraction error, i.e., the error between the true non-seasonal and its estimate. However, this raises another question scantly discussed in the literature. A time series model for $y_t^{tc} + y_t^i$, embedded in the seasonal adjustment procedure, namely used to predict future raw data, and a time series model for e_t, (properties often known and determined by the extraction filter), implies a model for $y_t^{tc} + y_t^i + e_t$. To the best of our knowledge applied time series studies never follow a strategy that borrows the non-seasonal component model used by statistical agencies and adds the stochastic properties of the extraction error to determine the prediction model for the seasonally adjusted series. Consequently, the model specification by statistical agencies in the course of seasonal adjusting a series is never taken into account when the adjusted series are actually used in forecasting exercises. Hence, seasonal adjustment and forecasting seasonally adjusted series are completely independent. In principle this ought not to be the case.

To conclude this subsection, it should be noted, however, that in some circumstances the filtering procedure is irrelevant and therefore the issues discussed in the previous paragraph are also irrelevant. The context is that of linear regression models with linear (seasonal adjustment) filters. This setting was originally studied by Sims (1974) and Wallis (1974), who considered regression models without lagged dependent variables; i.e., the classical regression. They showed that OLS estimators are consistent whenever all the series are filtered by the same filter. Hence, if all the series are adjusted by, say the linear X-11 filter, then there are no biases resulting from filtering. Absence of bias implies that point forecasts will not be affected by filtering, when such forecasts are based on regression models. In other words, the filter design is irrelevant as long as the same filter is used across all series. However, although parameter estimates remain asymptotically unbiased, it should be noted that residuals feature autocorrelation induced by filtering. The presence of autocorrelation should in principle be taken into account in terms of forecasting. In this sense, despite the invariance of OLS estimation to linear filtering, we should note that there remains an issue of residual autocorrelation.

5.3. Seasonal adjustment and feedback

While the topic of this Handbook is 'forecasting', it should be noted that in many circumstances, economic forecasts feed back into decisions and affect future outcomes. This is a situation of 'control', rather than 'forecasting', since the prediction needs to take into account its effect on future outcomes. Very little is said about the topic in this Handbook, and we would like to conclude this chapter with a discussion of the topic in the context of seasonal adjustment. The material draws on Ghysels (1987), who studies seasonal extraction in the presence of feedback in the context of monetary policy.

Monetary authorities often target nonseasonal components of economic time series, and for illustrative purpose Ghysels (1987) considers the case of monetary aggregates being targeted. A policy aimed at controlling the nonseasonal component of a time series can be studied as a linear quadratic optimal control problem in which observations are contaminated by seasonal noise (recall Equation (79)). The usual seasonal adjustment procedures assume however, that the future outcomes of the nonseasonal component are unaffected by today's monetary policy decisions. This is the typical forecasting situation discussed in the previous subsections. Statistical agencies compute future forecasts of raw series in order to seasonally adjusted economic time series. The latter are then used by policy makers, whose actions affect future outcomes. Hence, from a control point of view, one cannot separate the policy decision from the filtering problem, in this case the seasonal adjustment filter.

The optimal filter derived by Ghysels (1987) in the context of a monetary policy example is very different from X-11 or any of the other standard adjustment procedures. This implies that the use of (1) a model-based approach, as in SEATS/TRAMO, (2) a X-11-ARIMA or X-12-ARIMA procedure is suboptimal. In fact, the decomposition emerging from a linear quadratic control model is nonorthogonal because of the feedback. The traditional seasonal adjustment procedure start from an orthogonal decomposition. Note that the dependence across seasonal and nonseasonal components is in part determined by the monetary policy rule. The degree to which traditional adjustment procedures fall short of being optimal is difficult to judge [see, however, Ghysels (1987), for further discussion].

6. Conclusion

In this chapter, we present a comprehensive overview of models and approaches that have been used in the literature to account for seasonal (periodic) patterns in economic and financial data, relevant to forecasting context. We group seasonal time series models into four categories: conventional univariate linear (deterministic and ARMA) models, seasonal cointegration, periodic models and other specifications. Each is discussed in a separate section. A final substantive section is devoted to forecasting and seasonal adjustment.

The ordering of our discussion is based on the popularity of the methods presented, starting with the ones most frequently used in the literature and ending with recently proposed methods that are yet to achieve wide usage. It is also obvious that methods based on nonlinear models or examining seasonality in high frequency financial series generally require more observations than the simpler methods discussed earlier.

Our discussion above does not attempt to provide general advice to a user as to what method(s) should be used in practice. Ultimately, the choice of method is data-driven and depends on the context under analysis. However, two general points arise from our discussion that are relevant to this issue.

Firstly, the length of available data will influence the choice of method. Indeed, the relative lack of success to date of periodic models in forecasting may be due to the number of parameters that (in an unrestricted form) they can require. Indeed, simple deterministic (dummy variable) models may, in many situations, take account of the sufficient important features of seasonality for practical forecasting purposes.

Secondly, however, we would like to emphasize that the seasonal properties of the specific series under analysis is a crucial factor to be considered. Indeed, our Monte Carlo analysis in Section 2 establishes that correctly accounting for the nature of seasonality can improve forecast performance. Therefore, testing of the data should be undertaken prior to forecasting. In our context, such tests include seasonal unit root tests and tests for periodic parameter variation. Although commonly ignored, we also recommend extending these tests to consider seasonality in variance. If sufficient data are available, tests for nonlinearity might also be undertaken. While we are skeptical that nonlinear seasonal models will yield substantial improvements to forecast accuracy for economic time series at the present time, high frequency financial time series may offer scope for such improvements.

It is clear that further research to assess the relevance of applying more complex models would offer new insights, particularly in the context of models discussed in Sections 3 and 4. Such models are typically designed to capture specific features of the data, and a forecaster needs to be able to assess both the importance of these features for the data under study and the likely impact of the additional complexity (including the number of parameters estimated) on forecast accuracy.

Developments on the interactions between seasonality and forecasting, in particular in the context of the nonlinear and volatility models discussed in Section 4, are important areas of work for future consideration. Indeed, as discussed in Section 5, such issues arise even when seasonally adjusted data are used for forecasting.

References

Abeysinghe, T. (1991). “Inappropriate use of seasonal dummies in regression”. Economic Letters 36, 175–179.

Abeysinghe, T. (1994). “Deterministic seasonal models and spurious regressions”. Journal of Econometrics 61, 259–272.

Ahn, S.K., Reinsel, G.C. (1994). “Estimation of partially non-stationary vector autoregressive models with seasonal behavior”. Journal of Econometrics 62, 317–350.

Andersen, T.G., Bollerslev, T. (1997). “Intraday periodicity and volatility persistence in financial markets”. Journal of Empirical Finance 4, 115–158.

Andersen, T.G., Bollerslev, T., Lange, S. (1999). “Forecasting financial market volatility: Sample frequency vis-a-vis forecast horizon”. Journal of Empirical Finance 6, 457–477.

Barsky, R.B., Miron, J.A. (1989). “The seasonal cycle and the business cycle”. Journal of Political Economy 97, 503–535.

Beaulieu, J.J., Mackie-Mason, J.K., Miron, J.A. (1992). “Why do countries and industries with large seasonal cycles also have large business cycles?”. Quarterly Journal of Economics 107, 621–656.

Beaulieu, J.J., Miron, J.A. (1992). "A cross country comparison of seasonal cycles and business cycles". Economic Journal 102, 772–788.

Beaulieu, J.J., Miron, J.A. (1993). "Seasonal unit roots in aggregate U.S. data". Journal of Econometrics 55, 305–328.

Beltratti, A., Morana, C. (1999). "Computing value-at-risk with high-frequency data". Journal of Empirical Finance 6, 431–455.

Birchenhall, C.R., Bladen-Hovell, R.C., Chui, A.P.L., Osborn, D.R., Smith, J.P. (1989). "A seasonal model of consumption". Economic Journal 99, 837–843.

Bloomfield, P., Hurd, H.L., Lund, R.B. (1994). "Periodic correlation in stratospheric ozone data". Journal of Time Series Analysis 15, 127–150.

Bobbitt, L., Otto, M.C. (1990). "Effects of forecasts on the revisions of seasonally adjusted values using the X-11 seasonal adjustment procedure". In: Proceedings of the Business and Economic Statistics Section. American Statistical Association, Alexandria, pp. 449–453.

Bollerslev, T., Ghysels, E. (1996). "Periodic autoregressive conditional heteroscedasticity". Journal of Business and Economic Statistics 14, 139–151.

Boswijk, H.P., Franses, P.H. (1995). "Periodic cointegration: Representation and inference". Review of Economics and Statistics 77, 436–454.

Boswijk, H.P., Franses, P.H. (1996). "Unit roots in periodic autoregressions". Journal of Time Series Analysis 17, 221–245.

Box, G.E.P., Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Box, G.E.P., Tiao, G.C. (1975). "Intervention analysis with applications to economic and environmental problems". Journal of the American Statistical Association 70, 70–79.

Breitung, J., Franses, P.H. (1998). "On Phillips–Perron type tests for seasonal unit roots". Econometric Theory 14, 200–221.

Brockwell, P.J., Davis, R.A. (1991). Time Series: Theory and Methods, second ed. Springer-Verlag, New York.

Burridge, P., Taylor, A.M.R. (2001). "On the properties of regression-based tests for seasonal unit roots in the presence of higher-order serial correlation". Journal of Business and Economic Statistics 19, 374–379.

Busetti, F., Harvey, A. (2003). "Seasonality tests". Journal of Business and Economic Statistics 21, 420–436.

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 and Economic Statistics 13, 237–252.

Cecchitti, S., Kashyap, A. (1996). "International cycles". European Economic Review 40, 331–360.

Christoffersen, P.F., Diebold, F.X. (1998). "Cointegration and long-horizon forecasting". Journal of Business and Economic Statistics 16, 450–458.

Clements, M.P., Hendry, D.F. (1993). "On the limitations of comparing mean square forecast errors". Journal of Forecasting 12, 617–637.

Clements, M.P., Smith, J. (1999). "A Monte Carlo study of the forecasting performance of empirical SETAR models". Journal of Applied Econometrics 14, 123–141.

Clements, M.P., Hendry, D.F. (1997). "An empirical study of seasonal unit roots in forecasting". International Journal of Forecasting 13, 341–356.

Dagum, E.B. (1980). "The X-11-ARIMA seasonal adjustment method". Report 12-564E, Statistics Canada, Ottawa.

Davidson, J.E.H., Hendry, D.F., Srba, F., Yeo, S. (1978). "Econometric modelling of the aggregate time series relationship between consumers' expenditure income in the United Kingdom". Economic Journal 88, 661–692.

del Barrio Castro, T., Osborn, D.R. (2004). "The consequences of seasonal adjustment for periodic autoregressive processes". Econometrics Journal 7, 307–321.

Dickey, D.A. (1993). "Discussion: Seasonal unit roots in aggregate U.S. data". Journal of Econometrics 55, 329–331.

Dickey, D.A., Fuller, W.A. (1979). "Distribution of the estimators for autoregressive time series with a unit root". Journal of the American Statistical Association 74, 427–431.
Dickey, D.A., Hasza, D.P., Fuller, W.A. (1984). "Testing for unit roots in seasonal time series". Journal of the American Statistical Association 79, 355–367.
Engle, R.F., Granger, C.W.J., Hallman, J.J. (1989). "Merging short- and long-run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting". Journal of Econometrics 40, 45–62.
Engle, R.F., Granger, C.W.J., Hylleberg, S., Lee, H.S. (1993). "Seasonal cointegration: The Japanese consumption function". Journal of Econometrics 55, 275–298.
Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., Chen, B.-C. (1998). "New capabilities and methods of the X-12-ARIMA seasonal-adjustment program". Journal of Business and Economic Statistics 16, 127–177 (with discussion).
Franses, P.H. (1991). "Seasonality, nonstationarity and the forecasting of monthly time series". International Journal of Forecasting 7, 199–208.
Franses, P.H. (1993). "A method to select between periodic cointegration and seasonal cointegration". Economics Letters 41, 7–10.
Franses, P.H. (1994). "A multivariate approach to modeling univariate seasonal time series". Journal of Econometrics 63, 133–151.
Franses, P.H. (1995). "A vector of quarters representation for bivariate time-series". Econometric Reviews 14, 55–63.
Franses, P.H. (1996). Periodicity and Stochastic Trends in Economic Time Series. Oxford University Press, Oxford.
Franses, P.H., Hylleberg, S., Lee, H.S. (1995). "Spurious deterministic seasonality". Economics Letters 48, 249–256.
Franses, P.H., Kunst, R.M. (1999). "On the role of seasonal intercepts in seasonal cointegration". Oxford Bulletin of Economics and Statistics 61, 409–434.
Franses, P.H., Paap, R. (1994). "Model selection in periodic autoregressions". Oxford Bulletin of Economics and Statistics 56, 421–439.
Franses, P.H., Paap, R. (2002). "Forecasting with periodic autoregressive time series models". In: Clements, M.P., Hendry, D.F. (Eds.), A Companion to Economic Forecasting. Basil Blackwell, Oxford, pp. 432–452, Chapter 19.
Franses, P.H., Romijn, G. (1993). "Periodic integration in quarterly UK macroeconomic variables". International Journal of Forecasting 9, 467–476.
Franses, P.H., van Dijk, D. (2005). "The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production". International Journal of Forecasting 21, 87–105.
Gallant, A.R. (1981). "On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form". Journal of Econometrics 15, 211–245.
Ghysels, E. (1987). "Seasonal extraction in the presence of feedback". Journal of the Business and Economic Statistics 5, 191–194. Reprinted in: Hylleberg, S. (Ed.), Modelling Seasonality. Oxford University Press, 1992, pp. 181–192.
Ghysels, E. (1988). "A study towards a dynamic theory of seasonality for economic time series". Journal of the American Statistical Association 83, 168–172. Reprinted in: Hylleberg, S. (Ed.), Modelling Seasonality. Oxford University Press, 1992, pp. 181–192.
Ghysels, E. (1991). "Are business cycle turning points uniformly distributed throughout the year?". Discussion Paper No. 3891, C.R.D.E., Université de Montréal.
Ghysels, E. (1993). "On scoring asymmetric periodic probability models of turning point forecasts". Journal of Forecasting 12, 227–238.
Ghysels, E. (1994a). "On the economics and econometrics of seasonality". In: Sims, C.A. (Ed.), Advances in Econometrics – Sixth World Congress. Cambridge University Press, Cambridge, pp. 257–316.
Ghysels, E. (1994b). "On the periodic structure of the business cycle". Journal of Business and Economic Statistics 12, 289–298.
Ghysels, E. (1997). "On seasonality and business cycle durations: A nonparametric investigation". Journal of Econometrics 79, 269–290.

Ghysels, E. (2000). "A time series model with periodic stochastic regime switching, Part I: Theory". Macroeconomic Dynamics 4, 467–486.

Ghysels, E., Bac, C., Chevet, J.-M. (2003). "A time series model with periodic stochastic regime switching, Part II: Applications to 16th and 17th Century grain prices". Macroeconomic Dynamics 5, 32–55.

Ghysels, E., Granger, C.W.J., Siklos, P. (1996). "Is seasonal adjustment a linear or nonlinear data filtering process?". Journal of Business and Economic Statistics 14, 374–386 (with discussion). Reprinted in: Newbold, P., Leybourne, S.J. (Eds.), Recent Developments in Time Series. Edward Elgar, 2003, and reprinted in: Essays in Econometrics: Collected Papers of Clive W.J. Granger, vol. 1. Cambridge University Press, 2001.

Ghysels, E., Hall, A., Lee, H.S. (1996). "On periodic structures and testing for seasonal unit roots". Journal of the American Statistical Association 91, 1551–1559.

Ghysels, E., Harvey, A., Renault, E. (1996). "Stochastic volatility". In: Statistical Methods in Finance. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, vol. 14. North-Holland, Amsterdam.

Ghysels, E., Lee, H.S., Noh, J. (1994). "Testing for unit roots in seasonal time series: Some theoretical extensions and a Monte Carlo investigation". Journal of Econometrics 62, 415–442.

Ghysels, E., Lee, H.S., Siklos, P.L. (1993). "On the (mis)specification of seasonality and its consequences: An empirical investigation with US data". Empirical Economics 18, 747–760.

Ghysels, E., Osborn, D.R. (2001). The Econometric Analysis of Seasonal Time Series. Cambridge University Press, Cambridge.

Ghysels, E., Perron, P. (1996). "The effect of linear filters on dynamic time series with structural change". Journal of Econometrics 70, 69–97.

Gladyshev, E.G. (1961). "Periodically correlated random sequences". Soviet Mathematics 2, 385–388.

Goméz, V., Maravall, A. (1996). "Programs TRAMO and SEATS, instructions for the user (beta version: September 1996)". Working Paper 9628, Bank of Spain.

Hamilton, J.D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle". Econometrica 57, 357–384.

Hamilton, J.D. (1990). "Analysis of time series subject to changes in regime". Journal of Econometrics 45, 39–70.

Hansen, L.P., Sargent, T.J. (1993). "Seasonality and approximation errors in rational expectations models". Journal of Econometrics 55, 21–56.

Harvey, A.C. (1993). Time Series Models. Harvester Wheatsheaf.

Harvey, A.C. (1994). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Harvey, A.C. (2006). "Unobserved components models". In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam. Chapter 7 in this volume.

Harvey, A.C., Ruiz, E., Shephard, N. (1994). "Multivariate stochastic variance models". Review of Economic Studies 61, 247–264.

Hassler, U., Rodrigues, P.M.M. (2004). "Residual-based tests against seasonal cointegration". Mimeo, Faculty of Economics, University of Algarve.

Herwartz, H. (1997). "Performance of periodic error correction models in forecasting consumption data". International Journal of Forecasting 13, 421–431.

Hylleberg, S. (1986). Seasonality in Regression. Academic Press.

Hylleberg, S. (1992). Modelling Seasonality. Oxford University Press.

Hylleberg, S. (1995). "Tests for seasonal unit roots: General to specific or specific to general?". Journal of Econometrics 69, 5–25.

Hylleberg, S., Jørgensen, C., Sørensen, N.K. (1993). "Seasonality in macroeconomic time series". Empirical Economics 18, 321–335.

Hylleberg, S., Engle, R.F., Granger, C.W.J., Yoo, B.S. (1990). "Seasonal integration and cointegration". Journal of Econometrics 44, 215–238.

Johansen, S. (1988). "Statistical analysis of cointegration vectors". Journal of Economic Dynamics and Control, 231–254.

Johansen, S., Schaumburg, E. (1999). "Likelihood analysis of seasonal cointegration". Journal of Econometrics 88, 301–339.

Jones, R.H., Brelsford, W.M. (1968). "Time series with periodic structure". Biometrika 54, 403–407.

Kawasaki, Y., Franses, P.H. (2004). "Do seasonal unit roots matter for forecasting monthly industrial production?". Journal of Forecasting 23, 77–88.

Kunst, R.M. (1993). "Seasonal cointegration in macroeconomic systems: Case studies for small and large European countries". Review of Economics and Statistics 78, 325–330.

Kunst, R.M., Franses, P.H. (1998). "The impact of seasonal constants on forecasting seasonally cointegrated time series". Journal of Forecasting 17, 109–124.

Lee, H.S. (1992). "Maximum likelihood inference on cointegration and seasonal cointegration". Journal of Econometrics 54, 1–49.

Lee, H.S., Siklos, P.L. (1997). "The role of seasonality in economic time series: Reinterpreting money-output causality in U.S. data". International Journal of Forecasting 13, 381–391.

Lin, J., Tsay, R. (1996). "Cointegration constraint and forecasting: An empirical examination". Journal of Applied Econometrics 11, 519–538.

Löf, M., Franses, P.H. (2001). "On forecasting cointegrated seasonal time series". International Journal of Forecasting 17, 607–621.

Löf, M., Lyhagen, J. (2002). "Forecasting performance of seasonal cointegration models". International Journal of Forecasting 18, 31–44.

Lopes, A.C.B.S. (1999). "Spurious deterministic seasonality and autocorrelation corrections with quarterly data: Further Monte Carlo results". Empirical Economics 24, 341–359.

Lund, R.B., Hurd, H., Bloomfield, P., Smith, R.L. (1995). "Climatological time series with periodic correlation". Journal of Climate 11, 2787–2809.

Lütkepohl, H. (1991). Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.

Lyhagen, J., Löf, M. (2003). "On seasonal error correction when the process include different numbers of unit roots". SSE/EFI Working Paper, Series in Economics and Finance, 0418.

Martens, M., Chang, Y., Taylor, S.J. (2002). "A comparison of seasonal adjustment methods when forecasting intraday volatility". Journal of Financial Research 2, 283–299.

Matas-Mir, A., Osborn, D.R. (2004). "Does seasonality change over the business cycle? An investigation using monthly industrial production series". European Economic Review 48, 1309–1332.

Mills, T.C., Mills, A.G. (1992). "Modelling the seasonal patterns in UK macroeconomic time series". Journal of the Royal Statistical Society, Series A 155, 61–75.

Miron, J.A. (1996). The Economics of Seasonal Cycles. MIT Press.

Ng, S., Perron, P. (1995). "Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag". Journal of American Statistical Society 90, 268–281.

Novales, A., Flores de Fruto, R. (1997). "Forecasting with periodic models: A comparison with time invariant coefficient models". International Journal of Forecasting 13, 393–405.

Osborn, D.R. (1988). "Seasonality and habit persistence in a life-cycle model of consumption". Journal of Applied Econometrics, 255–266. Reprinted in: Hylleberg, S. (Ed.), Modelling Seasonality. Oxford University Press, 1992, pp. 193–208.

Osborn, D.R. (1990). "A survey of seasonality in UK macroeconomic variables". International Journal of Forecasting 6, 327–336.

Osborn, D.R. (1991). "The implications of periodically varying coefficients for seasonal time-series processes". Journal of Econometrics 48, 373–384.

Osborn, D.R. (1993). "Discussion on seasonal cointegration: The Japanese consumption function". Journal of Econometrics 55, 299–303.

Osborn, D.R. (2002). "Unit-root versus deterministic representations of seasonality for forecasting". In: Clements, M.P., Hendry, D.F. (Eds.), A Companion to Economic Forecasting. Blackwell Publishers.

Osborn, D.R., Chui, A.P.L., Smith, J.P., Birchenhall, C.R. (1988). "Seasonality and the order of integration for consumption". Oxford Bulletin of Economics and Statistics 50, 361–377. Reprinted in: Hylleberg, S. (Ed.), Modelling Seasonality. Oxford University Press, 1992, pp. 449–466.

Osborn, D.R., Smith, J.P. (1989). “The performance of periodic autoregressive models in forecasting seasonal UK consumption”. Journal of Business and Economic Statistics 7, 1117–1127.
Otto, G., Wirjanto, T. (1990). “Seasonal unit root tests on Canadian macroeconomic time series”. Economics Letters 34, 117–120.
Paap, R., Franses, P.H. (1999). “On trends and constants in periodic autoregressions”. Econometric Reviews 18, 271–286.
Pagano, M. (1978). “On periodic and multiple autoregressions”. Annals of Statistics 6, 1310–1317.
Payne, R. (1996). “Announcement effects and seasonality in the intraday foreign exchange market”. Financial Markets Group Discussion Paper 238, London School of Economics.
Reimers, H.-E. (1997). “Seasonal cointegration analysis of German consumption function”. Empirical Economics 22, 205–231.
Rodrigues, P.M.M. (2000). “A note on the application of DF test to seasonal data”. Statistics and Probability Letters 47, 171–175.
Rodrigues, P.M.M. (2002). “On LM-type tests for seasonal unit roots in quarterly data”. Econometrics Journal 5, 176–195.
Rodrigues, P.M.M., Gouveia, P.M.D.C. (2004). “An application of PAR models for tourism forecasting”. Tourism Economics 10, 281–303.
Rodrigues, P.M.M., Osborn, D.R. (1999). “Performance of seasonal unit root tests for monthly data”. Journal of Applied Statistics 26, 985–1004.
Rodrigues, P.M.M., Taylor, A.M.R. (2004a). “Alternative estimators and unit root tests for seasonal autoregressive processes”. Journal of Econometrics 120, 35–73.
Rodrigues, P.M.M., Taylor, A.M.R. (2004b). “Efficient tests of the seasonal unit root hypothesis”. Working Paper, Department of Economics, European University Institute.
Sims, C.A. (1974). “Seasonality in regression”. Journal of the American Statistical Association 69, 618–626.
Smith, R.J., Taylor, A.M.R. (1998). “Additional critical values and asymptotic representations for seasonal unit root tests”. Journal of Econometrics 85, 269–288.
Smith, R.J., Taylor, A.M.R. (1999). “Likelihood ration tests for seasonal unit roots”. Journal of Time Series Analysis 20, 453–476.
Taylor, A.M.R. (1998). “Additional critical values and asymptotic representations for monthly seasonal unit root tests”. Journal of Time Series Analysis 19, 349–368.
Taylor, A.M.R. (2002). “Regression-based unit root tests with recursive mean adjustment for seasonal and nonseasonal time series”. Journal of Business and Economic Statistics 20, 269–281.
Taylor, A.M.R. (2003). “Robust stationarity tests in seasonal time series processes”. Journal of Business and Economic Statistics 21, 156–163.
Taylor, S.J., Xu, X. (1997). “The incremental volatility information in one million foreign exchange quotations”. Journal of Empirical Finance 4, 317–340.
Tiao, G.C., Grupe, M.R. (1980). “Hidden periodic autoregressive moving average models in time series data”. Biometrika 67, 365–373.
Troutman, B.M. (1979). “Some results in periodic autoregression”. Biometrika 66, 219–228.
Tsiakas, I. (2004a). “Periodic stochastic volatility and fat tails”. Working Paper, Warwick Business School, UK.
Tsiakas, I. (2004b). “Is seasonal heteroscedasticity real? An international perspective”. Working Paper, Warwick Business School, UK.
van Dijk, D., Strikholm, B., Terasvirta, T. (2003). “The effects of institutional and technological change and business cycle fluctuations on seasonal patterns in quarterly industrial production series”. Econometrics Journal 6, 79–98.
Wallis, K.F. (1974). “Seasonal adjustment and relations between variables”. Journal of the American Statistical Association 69, 18–32.
Wells, J.M. (1997). “Business cycles, seasonal cycles, and common trends”. Journal of Macroeconomics 19, 443–469.
Whittle, P. (1963). Prediction and Regulation. English University Press, London.
Young, A.H. (1968). “Linear approximation to the census and BLS seasonal adjustment methods”. Journal of the American Statistical Association 63, 445–471.