Macroeconomic Dynamics, **4**, 2000, 467–486. Printed in the United States of America.

TIME-SERIES MODEL WITH PERIODIC STOCHASTIC REGIME SWITCHING

Part I: Theory

ERIC GHYSELS
University of North Carolina
and
Centre Interuniversitaire de Recherche en Analyse des Organisations

We present a class of stochastic regime-switching models. The time-series models may have periodic transition probabilities and the drifts may be seasonal. In the latter case, the model exhibits seasonal dummy variation that may change with the regime. The processes entail nontrivial interactions between so-called business and seasonal cycles. We discuss the stochastic properties as well as their relationship with periodic ARMA processes. Estimation and testing are also discussed in detail.

Keywords: Markov Regime-Switching Models, Periodic ARMA Processes, Time-Series Models

1. INTRODUCTION

The economy may recover from a slump much easier when seasonals are at their peak. Bankruptcies tend to be postponed until after a shopping season. Financial panics and stock market crashes tend to cluster around the fall.[1] Is there any reason to believe that these statements are true? The question is, can we model such interdependencies in a simple fashion?

To accomplish this task, we first need to introduce a class of models hitherto not studied formally in time-series analysis. Indeed, we need to formulate a class of models that would allow us to estimate parametrically the nontrivial (nonlinear) interdependencies between the different patterns of growth fluctuations just described and to allow us to test for such interdependencies. Hence, this paper breaks with a long-established tradition alluded to earlier of viewing seasonal fluctuations as separate and orthogonal to all other movements in the economy. Two types of models are considered. The first one that we analyze generalizes the periodic regime-switching model described by Ghysels (1991). It has important connections

Conversations with René Garcia, John Geweke, Jim Hamilton, Adrian Pagan, Tom Sargent, Chris Sims, Doug Steigerwald, and Mark Watson have proven very helpful during the course of this work. We also thank two referees for their invaluable comments. Benoit Durocher provided excellent research assistance. Address correspondence to: Eric Ghysels, Department of Economics, University of North Carolina, Chapel Hill, NC 27599-3305, USA; e-mail: eghysels@unc.edu.

 1365-1005/00 $9.50

to and also builds further on the regime-switching model of Hamilton (1989) as well as the hidden periodic structures presented by Tiao and Grupe (1980), Osborn (1988), and Hansen and Sargent (1996). The second class of models that we consider is inspired by Diebold et al. (1994), Durland and McCurdy (1994), and Filardo (1994), among others, who considered regime-switching models with (stochastic) time-varying transition probabilities using logistic functions. Besides univariate models, we also consider multivariate ones. Empirical applications of the models proposed here appear in a paper by Ghysels et al. (1998) and a companion paper by Bac et al. (1998).

The structure of the paper is as follows: Section 2 deals with the stochastic process theory, estimation, and hypothesis testing; Section 3 covers theoretical results; and Section 4 concludes.

2. STOCHASTIC PROCESS THEORY OF PERIODIC MARKOV REGIME-SWITCHING MODELS

A general class of periodic Markov regime-switching models is presented in this section. Special cases of this class include the (aperiodic) regime-switching models considered by Hamilton (1988, 1989, 1990), Phillips (1991), Cecchetti and Lam (1992), Albert and Chib (1993), Diebold et al. (1994), Durland and McCurdy (1994), Filardo (1994), Kim (1994), McCulloch and Tsay (1994), Sichel (1994), among others, as well as the periodic Markovian regime-switching structure presented by Ghysels (1991, 1994b), which was used to investigate the nonuniformity of the distribution of the NBER business-cycle turning points. The discussion focuses first on a simplified example that presents some of the key features and elements of interest. The main purpose of this example is to appeal to intuition for presenting the basic insights, deferring all technical and formal discussion to a later section. Section 2.1 sets the scene, introducing the notations as well as the specific model, which is an AR(1) stochastic regime-switching model with a periodic Markov chain. Section 2.2 elaborates on the linear ARMA and linear periodic representation of the (nonlinear) stochastic regime-switching AR(1) model. Properties such as the periodic duration distribution and seasonal conditional heteroskedasticity are highlighted in Section 2.3. A general framework and characterization for the class of periodic Markov regime-switching models are presented in Section 2.4.

2.1. A Univariate AR(1) Model as an Example

The purpose of this section is to provide motivation and insights by first using a simple model. Consider a univariate time-series process denoted $\{y_t\}$. It will typically represent a growth rate of, for example, GNP. Furthermore, let $\{y_t\}$ be generated by the following stochastic structure:

$$\{y_t - \mu[(i_t, \mathbf{s}_t)]\} = \phi\{y_{t-1} - \mu[(i_{t-1}, \mathbf{s}_{t-1})]\} + \varepsilon_t \tag{1}$$

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

with

$$\mu[(i_t, \mathbf{s}_t)] = \alpha_0 + \alpha_1 i_t + \sum_{s=1}^{\mathcal{S}-1} \mathbf{1}_{st}\alpha_s \tag{2}$$

where $|\phi| < 1$, ε_t is i.i.d. $N(0, \sigma^2)$ and $\mu[\cdot]$ represents an intercept shift function that includes seasonal dummies $\mathbf{1}_{st}$. If $\mu[(i_t, \mathbf{s}_t)] = \alpha_0 + \sum_{s=1}^{\mathcal{S}-1} \mathbf{1}_{st}\, \alpha_s$ with $\mathbf{s}_t = t \bmod(\mathcal{S})$, where $\mathcal{S}$ is the frequency of sampling throughout the year (e.g., $\mathcal{S} = 4$ for quarterly sampling), then (1) would simply be a standard linear stationary Gaussian AR(1) model with seasonal mean shifts α_s for $s = 1, \ldots, \mathcal{S}$. Instead, we assume that the intercept changes according to a Markovian regime-switching model, following the work of Hamilton (1989). The "state-of-the-world" process is different, however, from that originally considered by Hamilton. The state of the world is described by $(i_t, \mathbf{s}_t)$, which is a stochastic regime-switching process $\{i_t\}$ and the seasonal indicator process. Assuming that $i_t \in \{0, 1\}\ \forall\, t$, we allow the $\{i_t\}$ and $\{\mathbf{s}_t\}$ processes to interact in the following way[2]:

$$\begin{array}{ccc} & 0 & 1, \\ 0 & q(\mathbf{s}_t) & 1 - q(\mathbf{s}_t), \\ 1 & 1 - p(\mathbf{s}_t) & p(\mathbf{s}_t), \end{array} \tag{3}$$

where the transition probabilities $q(\cdot)$ and $p(\cdot)$ are allowed to change with $\mathbf{s}_t$, that is, the season. Because $\mathbf{s}_t$ is a mod $\mathcal{S}$ series, there are, of course, at most, $\mathcal{S}$ values for $q(\cdot)$ and $p(\cdot)$; that is, $q(\mathbf{s}_t) \in \{q^1, \ldots, q^{\mathcal{S}}\}$ and $p(\mathbf{s}_t) \in \{p^1, \ldots, p^{\mathcal{S}}\}$, where $q(\mathbf{s}_t) = q^s$ and $p(\mathbf{s}_t) = p^s$ for $\mathbf{s} = \mathbf{s}_t$. Naturally, when

$$p(\cdot) = \bar{p} \text{ and } q(\cdot) = \bar{q}, \tag{4}$$

we obtain the standard homogeneous Markov chain model considered by Hamilton. However, if for at least some $\mathbf{s}_t$ the transition probability matrix differs, we have a situation in which a regime shift will be more or less likely depending on the time of the year. Because $i_t \in \{0, 1\}$, the process $\{y_t\}$ has a mean shift α_0 in state 1 $(i_t = 0)$ and $\alpha_0 + \alpha_1$ in state 2. Equations (1) through (4) are a version of Hamilton's model with a periodic stochastic switching process. If we impose the condition that $\alpha_1 > 0$ state 1 with low mean drift is called a recession and state 2 an expansion, then, according to (3), we stay in a recession or move to an expansion with a probability scheme that depends on the season.

2.2. Linear ARMA and Periodic ARMA Representations of a Periodic Markov Regime-Switching Process

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 regime-switching structure, as represented by (1) through (4), and the more conventional linear ARMA processes as well as the periodic ARMA models discussed by Tiao and Grupe (1980),

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

Todd (1983, 1990), Osborn (1988), Osborn and Smith (1989), and Hansen and Sargent (1996), among others. Let us start by briefly explaining intuitively what drives the connections between the different models. The model described in Section 2.1, with y_t typically representing a growth series, is covariance stationary under suitable regularity conditions discussed later. Consequently, the process has a linear Wold MA representation. Yet, the time-series model presented in the preceding section provides a relatively parsimonious structure that determines nonlinearly predictable MA innovations. In fact, there are two layers beneath the Wold MA representation. One layer relates to *hidden periodicities*, as described by Tiao and Grupe (1980) or Hansen and Sargent (1996), 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 regime-switching model imposes *further structure* even after the hidden periodicities are uncovered. Indeed, there is a second layer that makes the innovations of the augmented system nonlinearly predictable. Hence, the model described in the preceding section also has nonlinearly predictable innovations as well as features of hidden periodicities.

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

$$i_t = [1 - q(\boldsymbol{s}_t)] + \lambda(\boldsymbol{s}_t) i_{t-1} + v_t(\boldsymbol{s}_t), \tag{5}$$

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

$$v_t(\boldsymbol{s}_t) = \begin{cases} [1 - p(\boldsymbol{s}_t)] & \text{with probability} \quad p(\boldsymbol{s}_t) \\ -p(\boldsymbol{s}_t) & \text{with probability} \quad 1 - p(\boldsymbol{s}_t), \end{cases} \tag{6}$$

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

$$v_t(\boldsymbol{s}_t) = \begin{cases} -[1 - q(\boldsymbol{s}_t)] & \text{with probability} \quad q(\boldsymbol{s}_t) \\ q(\boldsymbol{s}_t) & \text{with probability} \quad 1 - q(\boldsymbol{s}_t). \end{cases} \tag{7}$$

Equation (5) is a periodic AR(1) model where all of 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 $\{i_t\}$ but also by $\{\boldsymbol{s}_t\}$.[3] Although (5) resembles the periodic ARMA models that were discussed by Tiao and Grupe (1980), Todd (1983, 1990), Osborn (1988), and Hansen and Sargent (1996), among others, it is also fundamentally different in many respects. The most obvious difference is the innovation process, which has a discrete distribution. There are more subtle differences as well, but we shall highlight those as we further develop the model. Despite the differences, there are many features that equation (5) and the more standard periodic linear ARMA models have in common. Following Gladysev (1961), we can consider time-invariant

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

representations of (5) that are built on stacked, skip-sampled vectors of observations. In particular, let us assume that we have a sample of length $\mathcal{S}T$, that is, T number of years. Let us define the stacked vector of seasons that is sampled at an annual frequency as follows:

$$\boldsymbol{i}_\tau \equiv \left(i_{\mathcal{S}_\tau-\mathcal{S}+1}, i_{\mathcal{S}_\tau-\mathcal{S}+2}, \ldots \ldots, i_{\mathcal{S}_\tau}\right)', \tag{8}$$

$$\boldsymbol{v}_\tau \equiv \left(v^1_{\mathcal{S}_\tau-\mathcal{S}+1}, v^2_{\mathcal{S}_\tau-\mathcal{S}+2}, \ldots \ldots, v^{\mathcal{S}}_{\mathcal{S}_\tau}\right)', \tag{9}$$

$\tau = 1, \ldots, T$, so that τ represents annual time accounting. Following equation (5), we can write the DGP for the vector defined in (8) as follows:

$$\begin{bmatrix} 1 & 0 & \ldots & & 0 \\ -\lambda^2 & 1 & \ldots & & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & & & \\ 0 & & & 1 & 0 \\ 0 & \ldots & \ldots & -\lambda^{\mathcal{S}} & 1 \end{bmatrix} \boldsymbol{i}_\tau = \begin{bmatrix} 1-q^1 \\ 1-q^2 \\ \vdots \\ \\ \vdots \\ 1-q^{\mathcal{S}} \end{bmatrix} + \begin{bmatrix} 0 & \ldots & \lambda^1 \\ 0 & & 0 \\ & & \\ & & \\ 0 & \ldots & 0 \end{bmatrix} \boldsymbol{i}_{\tau-1} + \boldsymbol{v}_\tau . \tag{10}$$

We highlight two features of (10) on which we will digress further. The first is the appearance of seasonal mean shifts, that is, what typically is called "deterministic seasonality"; the second is the basis of a time-invariant Wold MA representation for the (scalar) $\{i_t\}$ process described by (5). We focus first on the latter, and follow with a discussion of the former.

The purpose of stacking the process $\{i_t\}$ into annual vectors is to exhaust all possible parameter variation appearing in (5) and (6). It is easy to see that the vector process $\{\boldsymbol{i}_\tau\}$ has a covariance stationary representation now (again under suitable regularity conditions discussed later) because the coefficient matrices in (10) are time invariant. Through (10), we can derive the Wold representation of $\{i_t\}$. This usually is referred to as the Tiao–Grupe formula. Because we will be explicitly using this formula, we briefly discuss it.[4] Assume that the Wold infinite-order MA decomposition representation for the vector process $\{\boldsymbol{i}_\tau\}$ can be written as follows:

$$\boldsymbol{i}_\tau = \boldsymbol{M}(L)\boldsymbol{\omega}_\tau + \boldsymbol{\mu}, \tag{11}$$

where $\boldsymbol{\omega}_\tau = [\omega_{\mathcal{S}(\tau-1)+1} \;\; \ldots \;\; \ldots \;\; \omega_{\mathcal{S}_\tau}]'$, $\boldsymbol{\mu} = (\mu_1 \;\; \ldots \;\; \ldots \;\; \mu_{\mathcal{S}})'$. Then, the covariance-generating function for the $(\boldsymbol{i}_\tau - \boldsymbol{\mu})$ process is defined as

$$S_{\boldsymbol{i}}(z) = (2\pi)^{-1}\boldsymbol{M}(z)\Omega\boldsymbol{M}(z^{-1})'. \tag{12}$$

From the covariance-generating function of the vector process $\{\boldsymbol{i}_\tau\}$, we can obtain the covariance-generating function for the scalar stochastic regime-switching process $\{i_t\}$, by using the Tiao–Grupe formula:

$$s_i(z) = Q(z)S_{\boldsymbol{i}}(z^{\mathcal{S}})Q(z^{-1})', \tag{13}$$

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

where $Q(z) = \mathcal{S}^{-1/2}[1 \;\; z \;\; \dots \;\; z^{\mathcal{S}-1}]$. One can calculate a spectral representation of the $\{i_t\}$ process from (13) or derive the linear time-domain representation.

Note that the process certainly will not be represented by an AR(1) process because part of the state space is "missing." A more formal argument can be derived directly from the analyses by Tiao and Grupe (1980) and Osborn (1991).[5] The periodic nature of autoregressive coefficients pushes the seasonality into annual lags of the AR polynomial emerging from (13) and substantially complicates the MA component.

Ultimately, of course, we are interested in the time-series properties of $\{y_t\}$ as it is generated by (1) through (4) and how its properties relate to linear ARMA and periodic ARMA representations of the same process. Because

$$y_t = \alpha_0 + \alpha_1 i_t + \sum_{s=1}^{\mathcal{S}-1} 1_{st}\alpha_s (1 - \phi L)^{-1}\varepsilon_t, \tag{14}$$

and ε_t was assumed to be Gaussian and independent, we can simply view $\{y_t\} - \sum_{s=1}^{\mathcal{S}-1} 1_{st}\alpha_s$ as the sum of two independent unobserved processes: $\{i_t\}$ and the process $(1-\phi L)^{-1}\varepsilon_t$. Clearly, all of the features just described about the $\{i_t\}$ process will be translated into similar features inherited by the observed process y_t, while $y_t - \sum_{s=1}^{\mathcal{S}-1} 1_{st}\alpha_s$ has the following linear time-series representation:

$$s_y(z) = \alpha_1^2 s_i(z) + \{1/[(1-\phi z)(1-\phi z^{-1})]\}(\sigma^2/2\pi). \tag{15}$$

This linear representation has hidden periodic properties which can be derived from (12) and a stacked skip-sampled version of the $(1-\phi L)^{-1}\varepsilon_t$ process. Finally, the vector representation obtained as such would inherit the nonlinear predictable features of $\{i_t\}$.

Let us briefly return to (11), or, alternatively, to (10). We observe that the linear representation has seasonal mean shifts that would appear as a "deterministic seasonal" in the univariate representation of y_t. Hence, besides the spectral density properties appearing in (15), which may or may not show peaks at the seasonal frequency, periodic Markov switching produces seasonal mean shifts in the univariate representation. This result is, of course, quite interesting because intrinsically we do 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. Note also, of course, that seasonal mean shifts appear in equation (2) already. Consequently, these mean-shift coefficients α_s in (2) are expected to differ from seasonal dummies appearing in a linear representation of $\{y_t\}$ that does not contain the Markov switching component.

2.3. Some Properties of Interest

So far, we have established some of the characteristics of the stochastic regime-switching AR(1) process with periodic transition probabilities. In particular, in the

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

preceding section, we described how to obtain a linear time-series representation and how it entails hidden periodicity and nonlinear predictability. In this section, we further digress on three properties of special interest: (1) seasonal conditional asymmetries, (2) the periodic duration distribution, and (3) the seasonal impulse response functions. We discuss each of these separately.

Seasonal conditional variance asymmetries. Consider the conditional variance of the innovation process appearing in (5). It can be written as

$$E\left[\left(v_t^s\right)^2 \,\middle|\, i_{t-1}, \boldsymbol{s}_t\right] = \begin{cases} p(\boldsymbol{s}_t)[1 - p(\boldsymbol{s}_t)] & \text{if } i_{t-1} = 1 \\ q(\boldsymbol{s}_t)[1 - q(\boldsymbol{s}_t)] & \text{if } i_{t-1} = 0. \end{cases} \qquad \textbf{(16)}$$

We observe that the variance of the stochastic regime-switching process, whether it is presented as a scalar or a vector, displays heteroskedasticity, conditional not only with regard to the season but also the regime shifts. The former source of heteroskedasticity, namely the seasonal variation in (conditional) second moments, is a natural by-product of the hidden periodicity and also features periodic ARMA processes. However, what is different is the asymmetry in conditional second moments blended with the periodic structure.

Periodic duration distribution. This feature highlights a characteristic proper to periodic Markov chains that was exploited by Ghysels (1991, 1997) to test the presence of periodicity via exact small-sample rank-based nonparametric statistics. If a Markov chain is periodic, then the distribution of the length of time spent in any particular regime depends on the starting season.

Seasonal impulse response functions. The purpose here is only to point out that, due to the hidden periodicity, there is also a periodic impulse response scheme that goes with the Wold decompositions conditional on the season as presented in (13). Hansen and Sargent (1996) studied in detail how the impulse response mechanisms operate in a periodic (linear) environment. We can only refer the reader to their detailed exposition. We also emphasize that Hansen and Sargent provided several examples of economic structural models that yield a linear periodic representation. Similar attempts were made by Todd (1983, 1990) and Osborn (1988), though Hansen and Sargent provided a unifying general equilibrium approach.

3. A GENERAL CLASS OF PERIODIC MARKOV STOCHASTIC REGIME-SWITCHING MODELS

Having mostly relied on intuition and on specific examples so far, we now turn our attention to generalizations. Here, we will only point to the different directions in which one can generalize the model and discuss how they can be formally treated. In Section 3.1, we present the general class of models and, in Section 3.2, we discuss the formal regularity conditions. In Section 3.3, we cover estimation and testing.

3.1. Model Structure

Consider the set $\mathcal{Y}$ of $\mathbf{R}^n$-valued discrete time vector processes defined on the probability space $(\Omega, \mathcal{S}, \mathcal{P})$, where for each $\omega \in \Omega$, $\{y_t(\omega)\} \in \mathcal{Y}$ is generated as follows:[6]

$$y_t = b_0(i_t, z_t)x_{0t} + \sum_{j=1}^{\ell} b_j(i_t, \mathbf{s}_t, z_t)[y_{t-j} - b_0(i_{t-j}, z_{t-j})x_{0t-j}] + \delta_t. \quad \mathbf{(17)}$$

Equation (17) is an explicit representation of the vector process, showing that possibly all coefficient matrices $b_j(\cdot)$, $j = 0, 1, \ldots, \ell$ are random and depend on the state process $(i_t, \mathbf{s}_t, z_t)$, where $\{i_t\}$ follows a Markov chain with transition probability matrix $P(\mathbf{s}_t, z_t)$ and $\mathbf{s}_t \equiv t \bmod \mathcal{S}$ as defined earlier, while z_t is a set of variables affecting the transition probabilities in a manner similar to that of Diebold et al. (1994), Durland and McCurdy (1994), and Filardo (1994), among others. The regressors x_{0t} in equation (17) are fixed, consisting of either a constant or a constant and $\mathcal{S} - 1$ seasonal dummies, while the error process δ_t is i.i.d. $N[0, \Lambda(i_t, \mathbf{s}_t, z_t)]$. Hence, the innovation variance may depend on the discrete state-of-the-world process.

A brief digression on the Markov process $\{i_t\}$ will be helpful before discussing the matrix functions $b_j(\cdot)$, $j = 0, \ldots, \ell$. It will be assumed that there are r "primitive" states describing r possible regimes. Because there are ℓ lags in equation (17), the Markov process will have $r^{\ell+1}$ states *each season*. Hence, the Markov chain throughout the year is described by the set $\{P(\mathbf{s}, z_t), \mathbf{s} = 1, \ldots, \mathcal{S}\}$, where $P(\mathbf{s}, z_t)$ is an $r^{\ell+1} \times r^{\ell+1}$ transition probability matrix. Following Diebold et al. (1994), Durland and McCurdy (1994), and Filardo (1994), we can consider the transition probabilities to be time varying, evolving as logistic functions of $z_t'\gamma_i(\mathbf{s})$, $\mathbf{s} = 1, \ldots, \mathcal{S}$. Hence, in state i, a different vector $\boldsymbol{\gamma}(\cdot)$ applies to each season. To illustrate this further, just let $\ell = 0$ and $r = 2$ for the moment. Then we have

$$\begin{array}{ccc} & 0 & 1 \\ 0 & \dfrac{\exp[z_t'\gamma_0(\mathbf{s})]}{\{1 + \exp[z_t'\gamma_0(\mathbf{s})]\}} & 1 - \dfrac{\exp[z_t'\gamma_0(\mathbf{s})]}{\{1 + \exp[z_t'\gamma_0(\mathbf{s})]\}} \\ 1 & 1 - \dfrac{\exp[z_t'\gamma_1(\mathbf{s})]}{\{1 + \exp[z_t'\gamma_1(\mathbf{s})]\}} & \dfrac{\exp[z_t'\gamma_1(\mathbf{s})]}{\{1 + \exp[z_t'\gamma_1(\mathbf{s})]\}}. \end{array} \quad \mathbf{(18)}$$

A special case of (18) is where z_t is a constant. Then the transition matrix simply becomes a function of $\mathbf{s}_t$ only, which is what appeared in (2). Because $\gamma_i(\cdot)$ becomes a scalar in such a case, we can simply express the transition probabilities $p(\cdot)$ and $q(\cdot)$ in (2) via the logistic function $q(\mathbf{s}) = \exp[\gamma_0(\mathbf{s})]/\{1 + \exp[\gamma_0(\mathbf{s})]\}$ and $p(\mathbf{s}) = \exp[\gamma_1(\mathbf{s})]/\{1 + \exp[\gamma_1(\mathbf{s})]\}$. Another special case of (18) that is also of particular interest for empirical work is when $P(\mathbf{s}_t, z_t) \equiv P(z_t)$, that is, independent of $\boldsymbol{s}_t$. Of course, this corresponds exactly to the analysis by Diebold et al. (1994). Yet, unlike the work of Diebold et al. (1994), Durland and McCurdy (1994), or Filardo (1994), it is important to observe that z_t may contain seasonal processes. Hence, the transition probability matrix becomes seasonal as well as time varying.

The $n \times n$ [where $n = \dim(y_t)$] matrix functions $b_1(\cdot)$ through $b_l(\cdot)$ appearing in (17) are allowed to shift with the regime. It was noted that the set of regressions x_{0t} consisted of a constant with or without seasonal dummies. With only a constant in x_{0t}, that is, $x_{0t} = 1 \,\forall\, t$, and $b_0(\cdot)$ only depending on $\{i_t\}$, we recover the most familiar case in which $\{y_t\}$ is driven by a stochastic mean shift which is a function of a latent Markov process determining the regime switches. Because y_t is an $n \times 1$ vector process, $b_0(\cdot)$ determines an $n \times 1$ vector of mean shifts depending on $\{i_t\}$ for the joint multivariate process. When x_{0t} also includes seasonal dummies it is worth noting that the seasonal mean shifts are allowed to shift with the regime, a feature discussed by Canova and Ghysels (1994).

Obviously, equation (17) contains many features all at once, making it potentially a richly parameterized model that will be too demanding from which to infer most data sets. Because $b_1, \ldots, b_\ell(\cdot)$ and $b_0(\cdot)$ are allowed to depend on $\{i_t\}$ and $\{\boldsymbol{s}_t\}$, one can indeed produce some quite complex dynamics in polynomial lags, seasonals, and regimes.[7]

3.2. Regularity Conditions

So far, we have presented a vector stochastic regime-switching process with possibly seasonal transition properties, both periodic through $\boldsymbol{s}_t$ and possibly stochastic through z_t, with fixed regressors and an AR(ℓ) polynomial autoregressive structure. When is such a process stable? When does it have finite moments, like a well-defined covariance structure, for instance? Our formal treatment only covers the case in which z_t is a constant, hence the transition matrix $P(\cdot)$ is nonrandom, but possibly periodic.[8] It is shown that (periodic) Markov regime-switching processes can be treated as doubly stochastic vector AR(1) processes, using the terminology coined by Tjøstheim (1986). Our formalization relies on Tjøstheim (1990) and Karlsen (1990) to characterize the necessary conditions for weak stationarity. An autoregressive representation for i_t, similar to equation (5) but more general, is the most convenient representation for our purpose. Let us transcribe the definition of i_t using a slightly different notation: Consider the identity matrix of dimension $r^{\ell+1}$ and let the ith column be denoted by e_i for $i = 1, \ldots, r^{\ell+1}$. Then, ξ_t will represent the state of the world; that is, $i_t = i \Leftrightarrow \xi_t = e_i$. Similarly to stacking i_t over an entire year as in (8), we also can obtain

$$\boldsymbol{\xi}_\tau \equiv [\xi'_{\mathcal{S}(\tau-1)+1}, \ldots, \xi'_{\tau\mathcal{S}}]', \tag{19}$$

where $\boldsymbol{\xi}_\tau$ is an $\mathcal{S} \times (r^{\ell+1})$ vector containing $\mathcal{S}$ entries equal to 1. The process $\{\boldsymbol{\xi}_\tau\}$ will have a *homogeneous* vector autoregressive representation of order 1 because it corresponds to a Markov chain with a homogeneous transition probability matrix obtained from the set $\{P(\boldsymbol{s}), \boldsymbol{s} = 1, \ldots, \mathcal{S}\}$. More precisely,

$$\boldsymbol{\xi}_\tau = \begin{bmatrix} I & 0 & & 0 \\ -F_1 & I & & \vdots \\ & & \ddots & \\ & & -F_{\mathcal{S}-1} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \ldots & F_{\mathcal{S}} \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & & \vdots \\ 0 & & 0 \end{bmatrix} \boldsymbol{\xi}_{\tau-1} + \gamma_\tau, \tag{20}$$

where F_s is entirely determined by $P(\boldsymbol{s})$ for each $\boldsymbol{s} = 1, \ldots \mathcal{S}$, and γ_τ is uncorrelated with $\boldsymbol{\xi}_{\tau-i}$ for $i > 0$.[9] Note that to obtain a homogeneous Markov chain representation of a periodic regime-switching model with ℓ autoregressive lags, we need to consider an $(\mathcal{S} \times r^{\ell+1})$-state system.

We now define the, by now familiar, stacked skip-sampled versions of the series $\{y_t\}$ and $\{\delta_t\}$. Moreover, we also introduce the processes $\tilde{y}_t^\ell$, δ_t^ℓ, $\tilde{y}_t$ and their stacked counterparts:

$$\tilde{y}_t \equiv y_t - b_0(i_t, \boldsymbol{s}_t)x_{0t}, \tag{21}$$

$$\tilde{y}_t^\ell \equiv [\tilde{y}_t' \ldots \tilde{y}_{t-\ell+1}']', \tag{22}$$

$$\delta_t^\ell \equiv [\delta_t' \; O_{n(\ell-1)\times 1}']', \tag{23}$$

$$\tilde{\boldsymbol{y}}_\tau^\ell \equiv \left\{ \left[\tilde{y}_{\mathcal{S}(\tau-1)+1}^\ell\right]', \ldots, \left(\tilde{y}_{\tau\mathcal{S}}^\ell\right)' \right\}', \tag{24}$$

$$\boldsymbol{\delta}_\tau^\ell \equiv \left\{ \left[\delta_{\mathcal{S}(\tau-1)+1}^\ell\right]', \ldots, \left(\delta_{\tau\mathcal{S}}^\ell\right)' \right\}', \tag{25}$$

$$B_0^\ell(\xi_t, \boldsymbol{s}_t) \equiv \left\{ [b_0(i_t, \boldsymbol{s}_t)x_{0t}]' O_{n(\ell-1)\times 1}' \right\}', \tag{26}$$

$$\mathcal{B}_0^\ell(\boldsymbol{\xi}_t) \equiv \left\{ B_0^\ell[i_{\mathcal{S}(t-1)+1}, 1]' \ldots B_0^\ell(i_{\mathcal{S}t}, \mathcal{S})' \right\}', \tag{27}$$

where the latter two are stacked versions of the intercept process $b_0(i_t, \boldsymbol{s}_t)$. Finally, it is straightforward to define

$$y_t^\ell \equiv B_0^\ell(\xi_t, \boldsymbol{s}_t) + \tilde{y}_t^\ell, \tag{28}$$

$$\boldsymbol{y}_\tau^\ell \equiv \mathcal{B}_0^\ell(\xi_\tau) + \tilde{\boldsymbol{y}}_\tau^\ell. \tag{29}$$

From (17) and the processes defined in (21) through (27), we can characterize a doubly stochastic representation:

$$\tilde{\boldsymbol{y}}_\tau^\ell = \mathcal{B}^\ell(\boldsymbol{\xi}_\tau)\tilde{\boldsymbol{y}}_{\tau-1}^\ell + \boldsymbol{\delta}_\tau^\ell, \tag{30}$$

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

where

$$\mathcal{B}^{\ell}(\boldsymbol{\xi}_\tau) \equiv M_1^{-1} \begin{bmatrix} 0 & -B^{\ell}\left(\xi_{\mathcal{S}_\tau}, \mathcal{S}\right) \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}$$

and $-B^{\ell}(\xi_t, \mathbf{s})$ is a matrix $\ln \times \ln$ with the first row containing $b_1(i_t, s), \ldots,$ $b_l(i_t, s)$, stacked on top of the matrix $(I_n \otimes I_{n-1} 0_{(n-1)l \times n})$; the matrix M_1 contains $-B^{\ell}(\xi_{\mathcal{S}(\tau-1)+1}, 1)$ through $-B^{\ell}(\xi_{\mathcal{S}_\tau - 1}, \mathcal{S}-1)$ below the block diagonal structure containing $nl \times nl$ identity matrices; and $E\boldsymbol{\delta}_\tau^{\ell}(\boldsymbol{\delta}_\tau^{\ell})' = I_{\mathcal{S}} \otimes \Lambda^{\ell}$, where $\Lambda^{\ell} = E\delta_\tau^{\ell}(\delta_\tau^{\ell})'$.[10]

The regularity conditions for the existence of a well-defined autocovariance structure for a general periodic Markov regime-switching process can be presented now. Several formulas that characterize the autocovariance structure are introduced, with each entailing different computational operations. The structure is as follows: (1) basic assumptions are presented first, (2) the steady state of the Markov process is discussed next, and, (3) a theorem summarizes the main result.

Basic assumptions.

Assumption 1. The processes $\{\tilde{\boldsymbol{y}}_\tau\}$, $\{\boldsymbol{\delta}_\tau\}$, and $\{\boldsymbol{\xi}_\tau\}$ are defined on a common probability space $(\Omega, \mathcal{S}_{\mathcal{S}}, P_{\mathcal{S}})$. The process $\{\boldsymbol{\xi}_\tau\}$ is a Markov chain that is stationary and ergodic with a finite number of states defined on the state space S with dimension $\mathcal{S}r^{\ell+1}$. It has a transition matrix denoted by $\mathcal{P}$.

For convenience of notation, we denote $r^{\ell+1}$ as K so that the number of states in S equals $\mathcal{S}K$.[11] To proceed with the next assumption, let us define the sigma algebra: $\mathcal{S}_{\mathcal{S}y}^{\tau} \equiv \{\boldsymbol{y}_u, u \leq \tau\}$.

Assumption 2. The matrix functions $B^{\ell}(\cdot)$ and $\mathcal{B}^{\ell}(\cdot)$ appearing in (30) are of dimension $(n\ell) \times (n\ell)$; $(\mathcal{S}n\ell) \times (\mathcal{S}n\ell)$, respectively, and are measurable functions with respect to $\mathcal{S}_{\mathcal{S}y}^{\tau}$. Likewise, the matrix functions $B_0(\cdot)$ and $\mathcal{B}_0^{\ell}(\cdot)$ also are measurable with regard to the same sigma algebra.

Assumption 3. The process $\{\boldsymbol{\delta}_\tau\}$ is a martingale difference sequence with regard to $\mathcal{S}_{\mathcal{S}y}^{\tau}$ and $E\boldsymbol{\delta}_\tau\boldsymbol{\delta}_\tau' = I_{\mathcal{S}} \otimes \Lambda^{\ell} < \infty$.

Because the estimation will be likelihood-based, we shall assume that $\boldsymbol{\delta}_t$ is i.i.d. $N(0, I_{\mathcal{S}} \otimes \Lambda)$ [as we did in equation (17)]. Alternatively, the analysis in this section can be used to construct a generalized methods of moments estimator [cf. Hansen (1982)]. Finally, we also make the assumption that follows.

Assumption 4. The processes $\{\boldsymbol{\delta}_\tau\}$ and $\{\boldsymbol{\xi}_\tau\}$ are mutually independent.

Covariance structure. The basic question of interest is: Under what circumstances are $\{y_t\}$ and its derived processes integrable in quadratic mean? That is, $\{y_t(\omega)\}$ belongs to the usual Hilbert space $L^2(\Omega, \mathcal{S}, P)$ or $\{\tilde{\boldsymbol{y}}_\tau(\omega)\}$ belongs to $L^2(\Omega, \mathcal{S}_{\mathcal{S}}, P_{\mathcal{S}})$. Because all processes have a doubly stochastic representation,

https://doi.org/10.1017/S136510050001703X Published online by Cambridge University Press

we rely on the analysis by Karlsen (1990) to develop necessary conditions for the existence of a well-defined covariance structure. We begin with a discussion of the first moments. Hence, we are interested in the mean of $\{y_t\}$ as it appears in (17). This, of course, means that we want to analyze the cross product of the stochastic process $b_0(\cdot)$ and the fixed regressors x_{0t}. From Assumption 1, we know that the Markov chain process has a unique steady-state distribution. For each season $\boldsymbol{s} = 1, \ldots \ldots, \mathcal{S}$, we characterize the steady-state distribution as the solution $\pi_{\boldsymbol{s}}$ to

$$\pi_{\boldsymbol{s}} = F_{\boldsymbol{s}} \pi_{\mathrm{mod}(\boldsymbol{s}-1)} \quad \text{with} \quad \pi_{\boldsymbol{s}}' \, l = 1 \qquad \boldsymbol{s} = 1, \ldots, \mathcal{S}, \tag{31}$$

where the matrices $F_1, \ldots, F_{\mathcal{S}}$ are given in (20) and several methods can be used to compute $\pi_{\boldsymbol{s}}$ [see, e.g., Hamilton (1994, p. 3065)]. Moreover, the steady-state distribution π of the skip-sampled Markov chain ξ_τ can be obtained easily either by computing $\pi = F\pi$ or as follows:

$$\pi = \mathcal{S}^{-1}(\pi_1' \ldots \ldots \pi_{\mathcal{S}}')' \quad \text{with} \quad \pi' l = 1. \tag{32}$$

Consider now the seasonal sampling of the x_{0t} process, and let its limit be denoted $\bar{x}_0^{\boldsymbol{s}} \boldsymbol{s} = 1, \ldots, \mathcal{S}$. If x_{0t} is a constant, then of course $\bar{x}_0^{\boldsymbol{s}} \equiv \bar{x}_0 \, \forall \boldsymbol{s}$; if x_{0t} includes seasonal dummies, then $\bar{x}_0^{\boldsymbol{s}}$ represents a different $n \times 1$ vector each season. Likewise, if x_{0t} contains stochastic elements, then $\bar{x}_0^{\boldsymbol{s}}$ is a stochastic limit based on seasonal sampling. Let

$$b_0^K(\boldsymbol{s}) \equiv \left[b_0(1, \boldsymbol{s})\bar{x}_0^{\boldsymbol{s}}, \ldots, b_0(K, \boldsymbol{s})\bar{x}_0^{\boldsymbol{s}}\right] \qquad \boldsymbol{s} = 1, \ldots, \mathcal{S} \tag{33}$$

be the matrix of all K possible mean shifts each season $\boldsymbol{s}$. Then, the mean of y_t conditional on season $\boldsymbol{s}$ is expressed as

$$Ey_t \mid \boldsymbol{s} = b_0^K(\boldsymbol{s})\pi_{\boldsymbol{s}} \qquad \boldsymbol{s} = 1, \ldots, \mathcal{S}, \tag{34}$$

while the mean unconditional of $\boldsymbol{s}$ is $Ey_t \equiv \mathcal{S}^{-1} \sum_{\boldsymbol{s}=1}^{\mathcal{S}} Ey_t \mid \boldsymbol{s}$. Some special cases of (34) are worth pointing out. For instance, if x_{0t} is just a constant and $b_0^K(\boldsymbol{s})$ is not a function of $\boldsymbol{s}$, as for instance is the case in Hamilton (1989), then $Ey_t \mid \boldsymbol{s}$ is the cross product of $\pi_{\boldsymbol{s}}$ with b_0^K, that is, the expected mean shift under steady-state distribution $\pi_{\boldsymbol{s}}$. Hence, as observed in Section 2 for a specific case, with switching probabilities changing periodically, one generates a seasonal mean-shifting behavior in linear representation. This seasonal dummy behavior is tightly parameterized because it is determined entirely by the switching probabilities of the Markov chain. When x_{0t} includes dummies, a seasonal mean-shifting behavior naturally arises with a more flexible parameterization. Having determined the mean of y_t, conditional or unconditional on the season, we turn our attention next to the second moments of the demeaned process $\tilde{y}_t$ as specified in (23), as well as the second moments of the $b_0(i_t, \boldsymbol{s}_t)$ process. To streamline its characterization, we rely on the doubly stochastic representation appearing in (30). We are interested in the following objects:

$$\Gamma(H) = E y_\tau^\ell y_{\tau+H}^{\ell'}, \tag{35}$$

$$\Gamma_0(H) = E \mathcal{B}_0^\ell(\boldsymbol{\xi}_\tau) \mathcal{B}_0^\ell(\boldsymbol{\xi}_{\tau+H})', \tag{36}$$

$$\gamma_{\mathbf{s}}(h) = E y_t^\ell y_{t+h}^{\ell'} \, \forall\, t \text{ such that } t = \tau(\mathcal{S}-1) + \mathbf{s} \text{ and } h = 0, 1, \ldots, \tag{37}$$

$$\gamma_{0\mathbf{s}}(h) = E B_0^\ell(\xi_t, \mathbf{s}) B_0^\ell(\xi_{t+h}, \overline{\mathbf{s}})' \, \forall\, t \text{ such that } t = \tau(\mathcal{S}-1) + \mathbf{s}, \tag{38}$$

where $\tilde{\mathbf{s}} = (\mathbf{s} + h - 1) \bmod \mathcal{S}$ and $h = 0, 1, \ldots$.

The formula in (35) represents the covariance structure for the stacked skip-sampled vector process y_τ^ℓ. In contrast, the formula in (36) represents the covariance structure, conditional on a particular season, of the nonstacked y_t^ℓ process. Once the formulas in (35) and (36) are well defined and characterized, we can again invoke the Tiao–Grupe formula appearing in (13), this time applied to the y_t^ℓ process, yielding expressions for

$$\gamma(h) = E y_t^\ell y_{t+h}^{\ell\prime} \quad \forall t \text{ and } h = 0, 1, \ldots, \tag{39}$$

$$\gamma_0(h) = E B_0^\ell(\xi_t, \cdot) B_0^\ell(\xi_{t+h}, \cdot)' \quad \forall t \text{ and } h = 0, 1, \ldots. \tag{40}$$

The existence and characterization of (35) through (37) are determined as follows:

THEOREM 1. *Let Assumptions* 1 *through* 4 *hold. Then stochastic processes* $\{y_\tau^\ell\}$ *and* $\{B_0^\ell(\boldsymbol{\xi}_\tau)\}$ *are covariance-stationary with* $\Gamma(0)$ *and* $\Gamma_0(0)$ *finite if*

(*a*) $$\max_{1 \le \mathbf{s} \le \mathcal{S}} \left[\pi_{\mathbf{s}}' \otimes I_{(n\ell)^2}\right] \left\{\operatorname{diag}\left[B_0^\ell(k, \mathbf{s}) B_0^\ell(k, \mathbf{s})'\right]_{k=1}^K\right\} \left[l_K \otimes I_{(n\ell)^2}\right] < \infty,$$

(*b*) $$\rho\left\{\operatorname{diag}[\mathcal{B}^\ell(k) \otimes \mathcal{B}^\ell(k)']_{k=1}^{\mathcal{S}K}\right\}\left[\mathcal{P}' \otimes I_{(\mathcal{S}n\ell)^2}\right] < 1.$$

Finally,

$$\Gamma(H) = \begin{bmatrix} \gamma_1(\mathcal{S}) & \gamma_1(\mathcal{S}H - \mathcal{S} + 1) \\ \vdots & \\ \gamma_{\mathcal{S}}(\mathcal{S}H - \mathcal{S} + 1) & \gamma_{\mathcal{S}}(\mathcal{S}H) \end{bmatrix}, \tag{41}$$

with a similar relation applying to $\Gamma_0(H)$ *and* $\gamma_{0\mathbf{s}}^s(\cdot)$, *while*

$$\operatorname{vec}\gamma_{\mathbf{s}}(h) = \pi_{\tilde{\mathbf{s}}}' \otimes I_{(n\ell)^2}\left\{\operatorname{diag}\left[I_{\mathcal{S}n\ell} \otimes \mathcal{B}^\ell(k)\right]_{k=1}^{sK}\right\}\operatorname{vec}\gamma_{\mathbf{s}}(h-1) + \operatorname{vec}\Lambda^\ell. \tag{42}$$

The proof of the theorem is omitted here because it relies extensively on Theorem 4.1 of Karlsen (1990). By formulating equation (17) as a doubly stochastic vector AR(1) process, Karlsen (1990) spells out the conditions for a well-defined second-order structure when the parameter process in a doubly stochastic vector AR(1) process is governed by a finite Markov chain.

3.3. Estimation and Testing

Estimation of Markov regime-switching models is covered in detail by Hamilton (1989, 1994). Testing the hypothesis of periodicity in regime switching is of key interest for our purpose and is the focus of this section. Estimating a Markov regime-switching model is mostly likelihood-based, either via classical methods such as that used by Hamilton (1989) among others, or via Bayesian methods following the work of Albert and Chib (1993) and McCulloch and Tsay (1994). The estimation of Markov regime-switching models is covered in detail by Hamilton (1989, 1993) for the case in which the Markov chain is homogeneous and by Diebold et al. (1994) for time-varying transitions. Therefore, we restrict ourselves here to a brief discussion, only highlighting new features occurring because of periodicity. We first discuss the formulation of the likelihood function and then cover classical hypothesis testing.

In general, we seek to estimate the parameter vector $\boldsymbol{\Theta}$ governing the coefficient matrices $b_j(\cdot)$, $j = 0, 1, \ldots, \ell$, and the covariance matrix Λ from equation (17). We will make some concessions regarding generality and focus instead on the special case of two primitive states with a simple periodic Markov chain, such as in (3), involving a scalar stochastic process y_t, that is, $n = 1$. Given a sample of size $\mathcal{ST}$, that is, T full years of data points, the log-likelihood function can be written as

$$L(Y_{\mathcal{ST}}, \Theta) \equiv \sum_{t=1}^{\mathcal{ST}} \log p(y_t \mid Y_{t-1}; \boldsymbol{\Theta}), \tag{43}$$

where $p(\cdot \mid Y_{t-1}; \boldsymbol{\Theta})$ represents the probability distribution of y_t, given observations up to $t-1$; that is, $Y_{t-1} \equiv \{(y_{t-j}, \boldsymbol{s}_{t-j}),\ j \geq 1\}$. Hamilton (1989, 1993) goes into detail about how to formulate $p(\cdot \mid Y_{t-1}; \boldsymbol{\Theta})$ via a filtering algorithm to calculate the distribution of the time t state, given Y_t, denoted $p(\xi_t \mid Y_t)$, and future observations of y_t, in case of smoothed inference. The key element of interest is the probability of the unobservable state process $s_t = (\xi_t, \boldsymbol{s}_t)$ at any given point in time. This probability can be written as

$$p[s_t = (i_t, \boldsymbol{s}_t) \mid Y_j; \boldsymbol{\Theta}], \tag{44}$$

where j can be smaller than, equal to, or greater than t. The algorithm starts out with the unconditional probability. The first observation is drawn from $\pi_{\boldsymbol{s}}$ with $\boldsymbol{s} = \boldsymbol{s}_1$ and $\pi_{\boldsymbol{s}}$ being determined by (31). The unconditional probabilities for any of the seasons depend on all of the switching probabilities $p(\boldsymbol{s})$ and $q(\boldsymbol{s})$. Explicit formulas for the case $i_t \in \{0, 1\}$ appear in Ghysels [1994, eq. (2.6)]. Once $p[s_t = (1, \boldsymbol{s}_1)]$, for example, is computed, we can derive the joint probability of s_1 and s_2 as

$$p[s_2 = (1, \boldsymbol{s}_2), s_1 = (1, \boldsymbol{s}_1)] = p(\boldsymbol{s}_1)\, p[s_t = (1, \boldsymbol{s}_1)]. \tag{45}$$

Iterations similar to (47) can be computed for an initial segment $t = 1, \ldots, \ell + 1$, where ℓ is the lag length of the AR polynomial in (17). Then, the density of the $\ell + 1$ sample points conditional on $(s_1, \ldots, s_{\ell+1})$, which will be denoted

$p(Y_{\ell+1} \mid s_{\ell+1}, \ldots, s_1; \Theta)$, can be readily computed from the normality of δ_t in (17). The conditional distribution of the $\ell+1$, first states, given data points in $Y_{\ell+1}$, can be expressed as

$$p(s_{\ell+1}, \ldots, s_1 \mid Y_{\ell+1}) = \frac{p(Y_{\ell+1} \mid s_{\ell+1}, \ldots, s_1)\, p(s_{\ell+1}, \ldots, s_1)}{\sum_{k_1}^{K} \cdots \sum_{k_{\ell+1}}^{K} p(Y_{\ell+1} \mid s_{\ell+1} = k_{\ell+1}, \ldots, s_1 = k_1)\, p(s_{\ell+1} = k_{\ell+1}, \ldots, s_1 = k_1)}. \tag{46}$$

Equation (46) is applied iteratively throughout the sample, beginning with

$$p[s_{t+1} = (k, \boldsymbol{s}_{t+1}), s_t = (j, \boldsymbol{s}_t) \mid (s_{t-1}, \ldots, s_{t-\ell+1}) Y_t] = \left[\sum_{s=1}^{\mathcal{S}} \mathbf{1}\,(\boldsymbol{s}_{t+1} = \boldsymbol{s})\, p_{kj}^{(s)} \right] p[s_t \mid (s_{t-1}, \ldots, s_{t-\ell+1}) Y_t], \tag{47}$$

where $\mathbf{1}\,(\boldsymbol{s}_t = \boldsymbol{s})$ is a seasonal indicator function, namely,

$$\mathbf{1}\,(\boldsymbol{s}_t = \boldsymbol{s}) = \begin{cases} 1 & \text{if } \boldsymbol{s}_t = \boldsymbol{s} \\ 0 & \text{otherwise.} \end{cases}$$

Next, we can write

$$p[y_{t+1} \mid s_{t+1}, s_t, (s_{t-1}, \ldots, s_{t-\ell+1}) Y_{t+1}] \sim N\Bigg(b_0(i_{t+1}, \boldsymbol{s}_{t+1}) x_{0t+1} + \sum_{j=1}^{\ell} b_j(i_{t+1}, \boldsymbol{s}_{t+1}) \{[y_{t-j+1} - b_0(i_{t-j+1}, \boldsymbol{s}_{t-j+1})] x_{0t-j+1}\}, \Lambda \Bigg). \tag{48}$$

Combining equations (46) through (48) yields

$$p(s_{t+1} = (k, \boldsymbol{s}_{t+1}) \mid Y_{t+1}) = \sum_{j=1}^{K} \frac{p(y_{t+1} \mid s_{t+1}, s_t, Y_t) \left[\sum_{s=1}^{\mathcal{S}} \mathbf{1}\,(\boldsymbol{s}_{t+1} = \boldsymbol{s})\, p_{kj}(\boldsymbol{s}) \right] p[s_t = (j, \boldsymbol{s}_t) \mid Y_t]}{\left\{ \sum_{u=1}^{K} \sum_{v=1}^{K} p(y_{t+1} \mid s_{t+1}, s_t, Y_t) \left[\sum_{s=1}^{\mathcal{S}} \mathbf{1}\,(\boldsymbol{s}_{t+1} = \boldsymbol{s})\, p_{uv}(\boldsymbol{s}) \right] p[s_t = (v, \boldsymbol{s}_t) \mid Y_t] \right\}}. \tag{49}$$

The expressions in (50) together with (45) yield the desired log-likelihood function. One important feature about (50) will be most useful with respect to the derivation of an LM statistic: Only the $p(\boldsymbol{s})$ for $\boldsymbol{s}_{t+1} = \boldsymbol{s}$ appears in the recursion formula for $p(s_{t+1} \mid Y_{t+1})$. All other transition matrices $p(\boldsymbol{s})$ with $\boldsymbol{s} \neq \boldsymbol{s}_{t+1}$ do not appear directly, though, of course, they affect $p(s_t \mid Y_{t+1})$ on the right-hand side of (50).

Finally, we elaborate on the LM test discussed in this section, which consists of a system of $\mathcal{S}$ stacked score functions involving only transitions from a particular season. More specifically, in case there are two states, that is, $K=2$, we have for $i=1,2$ the following:

$$r_{it}^{s}(\lambda) \equiv \frac{\partial \log p(y_t \mid Y_{t-1};\lambda)}{\partial p_{ii}(\mathbf{s})} = \mathbf{1}(\mathbf{s}_t=\mathbf{s})\big\{p_{ii}(\mathbf{s})^{-1}p[s_t=(i,\mathbf{s}_t), s_{t-1}(i,\mathbf{s}_{t-1}) \mid Y_t]$$
$$-[1-p_{ii}(\mathbf{s})]^{-1}p[s_t=(j,\mathbf{s}_t), s_{t-1}=(i,\mathbf{s}_{t-1}) \mid Y_t]\big\}$$
$$+[p_{ii}(\mathbf{s})]^{-1}\left(\sum_{n=2}^{t-1}\mathbf{1}(\mathbf{s}_n=\mathbf{s})\{p[s_n=(i,\mathbf{s}_n), s_{n-1}=(i,\mathbf{s}_{n-1}) \mid Y_t]\right.$$
$$\left. - p(s_n=(i,\mathbf{s}_n), s_{n-1}=(i,\mathbf{s}_{n-1}) \mid Y_{t-1})\}\right)$$
$$-[1-p_{ii}(\mathbf{s})]^{-1}\left(\sum_{j=1}^{t-1}\mathbf{1}(\mathbf{s}_n=\mathbf{s})\{p[s_n=(j,\mathbf{s}_n), s_{n-1}=(i,\mathbf{s}_{n-1}) \mid Y_t]\right.$$
$$\left. - p[s_n=(j,\mathbf{s}_n), s_{n-1}=(i,\mathbf{s}_{n-1}) | Y_{t-1}]\}\right)$$
$$+\mathbf{1}(\mathbf{s}_1=\mathbf{s})\frac{p[s_1=(i,\mathbf{s}_1) \mid Y_t]-p[s_1=(i,\mathbf{s}_1) \mid Y_{t-1}]}{1-p_{ii}(\mathbf{s})} \quad \mathbf{(50)}$$

for $t \geq 2$, whereas for $t=1$,

$$r_{i1}^{s}(\lambda) = \mathbf{1}(\mathbf{s}_1=\mathbf{s})\frac{p[s_1=(i,\mathbf{s}_1) \mid Y_t]-\{[1-p_{jj}(\mathbf{s})]/[1-p_{ii}(\mathbf{s})+1-p_{jj}(\mathbf{s})]\}}{1-p_{ii}(\mathbf{s})}, \quad \mathbf{(51)}$$

where $j=2$ when $i=1$ and $j=1$ for $i=2$. From (51) and (52), we define

$$R_t^{s}(\lambda) \equiv \begin{bmatrix} r_{1t}^{s}(\lambda) \\ r_{2t}^{s}(\lambda) \end{bmatrix} \quad \mathbf{(52)}$$

and $R_t(\lambda) \equiv [R_t^1(\lambda)' \ldots R_t^{\mathcal{S}}(\lambda)']'$. The latter will be used to define the LM test.

Because general specification tests have been developed elsewhere, we will not devote much attention to their presentation. Indeed, Hamilton (1996) developed tests for omitted autocorrelation, omitted ARCH, and misspecification of the Markovian dynamics. Such tests can be applied easily to the present framework. In the remainder of the section, we focus our attention exclusively on the principal hypothesis of interest, namely, the periodicity of the Markov structure. The hypothesis of no periodic structure can be formally stated as follows:

$$H_0 : p_{ij}(\mathbf{s}) = \overline{p}_{ij} \quad \forall i, j \in \{1, \ldots, K\}, \mathbf{s} = 1, \ldots, \mathcal{S}. \tag{53}$$

This hypothesis is “standard,” and hence does not involve issues such as testing when nuisance parameters are not identified under the null and issues that emerge when testing model (17) against a linear time-series model, for instance.[12]

We first use an LR test, which can be formulated as follows:

$$LR = -2[L(y_{\mathcal{S}T}, \hat{\Theta}_c) - L(y_{\mathcal{S}T}, \hat{\Theta}_u)] \xrightarrow{d} \chi^2(\text{df}), \tag{54}$$

where $\hat{\Theta}_u$ and $\hat{\Theta}_c$ are the unrestricted and restricted ML estimates, respectively, df is the number of degrees of freedom equal to $(\mathcal{S} - 1) \times K$; and $L(\cdot)$ is the log likelihood appearing in (45) through (52).

Next, we consider an LM test for the same hypothesis. Such a test has several advantages over the LR test. Since one estimates the restricted (i.e., nonperiodic) model, one uses the score function of the periodic model evaluated at $\hat{\Theta}_c$. The fact that one has to estimate the model only once is one advantage. However, the most important advantage is that the parameter space is greatly reduced to the simple aperiodic model for which estimates are readily available, like the estimates obtained by Hamilton (1989) for the case of unadjusted quarterly GNP. The LM test is also elegant because of its structure. So far, we have shown that the conditional probability $p(s_{t+1} \mid Y_{t+1})$ only involves $p_{ij}(s)$ for $\mathbf{s}_{t+1} = \mathbf{s}$. Because of this feature the LM test will consist of a system of $\mathcal{S}$ stacked score functions involving only transitions from a particular season. More specifically, in case there are two states, that is, $K = 2$, we have for $i = 1, 2$ the following:

$$R_t^{\mathbf{s}}(\lambda) \equiv \begin{bmatrix} r_{1t}^{\mathbf{s}}(\lambda) \\ r_{2t}^{\mathbf{s}}(\lambda) \end{bmatrix},$$

where $r_{it}^{\mathbf{s}}(\lambda), i = 1, 2$, are defined in (50) and $R_t(\lambda) \equiv [R_t^1(\lambda)' \ldots R_t^{\mathcal{S}}(\lambda)']'$. The latter is a stacked system of score functions for each season. Then, the score test can be formulated as[13]

$$LM \equiv T \left[(1/T) \sum_{t=1}^{\mathcal{S}T} R_t(\hat{\lambda}_c) \right]' \left[(1/T) \sum_{t=1}^{\mathcal{S}T} R_t(\hat{\lambda}_c) R_t(\hat{\lambda}_c)' \right]^{-1}$$
$$\times \left[(1/T) \sum_{t=1}^{\mathcal{S}T} R_t(\hat{\lambda}_c) \right] \xrightarrow{d} \chi^2[(\mathcal{S} - 1) \times K]. \tag{55}$$

The statistic has a standard asymptotic distribution because the periodic structure is obviously an observable state.

4. CONCLUSIONS

We have presented a general class of regime-switching models that feature some intriguing properties that allow for some nontrivial interactions between seasonal cycles and so-called business cycles. Consequently, the regime-switching models introduced here break with a long tradition of viewing seasonal and business cycles as completely independent time-series movements. In a companion paper [Bac et al. (1998)] we provide an empirical application involving historical time series. Ghysels et al. (1998) also discuss several empirical examples applied to monthly and quarterly U.S. data series.

NOTES

1. See, for example, Shiller (1994) who counts 23 crashes since 1926, with 10 occurring in October.

2. To avoid notation that is too cumbersome, we do not introduce a separate notation for the theoretical representation of stochastic processes and their actual realizations.

3. If arguments $\boldsymbol{s}_t$ were absent from equations (5) and (6), then we would recover the nonperiodic AR(1) representation of Hamilton's (1989) stochastic regime-switching model as it appears in equations (3) and (4) of that paper.

4. For a more elaborate discussion, see Tiao and Grupe's (1980) original paper or Hansen and Sargent's (1996) paper.

5. Osborn (1991) in fact 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 by Ghysels (1993).

6. We substitute y_t for $y_t(\omega)$, etc., to avoid unnecessary notational complexity.

7. See, for instance, Hansen (1992) and McCulloch and Tsay (1994) for regime-switching models with state-dependent AR polynomials.

8. A formal treatment of such regularity conditions has been absent from the literature on Hamilton-type models. Because periodic Markov regime-switching models cover as a special case an aperiodic homogeneous Markov scheme, the regularity conditions apply to a large set of applications hitherto treated informally.

9. The correspondence between F_s and $P(\boldsymbol{s})$ is relatively simple and can be found, for instance, in Hamilton (1994, pp. 3063, 3067).

10. Notice that $\boldsymbol{s}_t$ no longer appears in $\mathcal{B}^{\ell}(\cdot)$ because it is absorbed through stacking.

11. The probability space used in Assumption 1 is appropriate for dealing with stacked skip-sampled vectors where stacking is based on seasons. In particular, $\mathcal{S}_S$ represents a sigma algebra based on sampling events conditional on the season in which they occur. The formal discussion presented here includes as special cases models that do not involve periodic Markov chains. Indeed, this is easy to see: Simply replace τ with t, and replace the probability space $(\Omega, \mathcal{S}_S, P_S)$ with $(\Omega, \mathcal{S}, P)$ while the number of states in Assumption 2 equals $r^{\ell+1}$. See Hansen and Sargent (1996) for the measure-theoretic issues involved.

12. Hansen (1991) discusses testing Hamilton's model against a linear time-series model. Using a standardized LR test, he was unable to reject the hypothesis of an AR(4) in favor of Hamilton's model. Instead, he found supporting evidence for a mixture model with a state-dependent AR(2) model.

13. Note that the normalization is with regard to T because the score test is based on a stacked vector central-limit argument.

REFERENCES

Albert, J. & S. Chib (1993) Bayesian inference via Gibbs sampling of autoregressive time series subject to Markov measured variance shifts. *Journal of Business and Economic Statistics* 11, 1–15.

Bac, C., J.-M. Chevet & E. Ghysels (in press) Time-series model with periodic stochastic regime switching. Part II: Applications to sixteenth and seventeenth century grain prices. *Macroeconomic Dynamics*.

Canova, F. & E. Ghysels (1994) Changes in seasonal patterns: Are they cyclical? *Journal of Economic Dynamics and Control* 18, 1143–1171.

Cecchetti, S.G. & P. Lam (1992) What Do We Learn from Variance Ratio Statistics? A Study of Stationary and Nonstationary Models with Breaking Trends. Mimeo, Ohio State University.

Diebold, F.X., J.-H. Lee & G.C. Weinbach (1994) Regime switching with time-varying transition probabilities. In C. Hargreaves (ed.), *Nonstationarity Time Series Analysis and Cointegration*. Oxford: Oxford University Press.

Durland, M.J. & T.H. McCurdy (1994) Duration-dependent transitions in a Markov model of U.S. GNP growth. *Journal of Business and Economic Statistics* 22, 279–288.

Filardo, A.J. (1994) Business cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics* 12, 299–308.

Ghysels, E. (1991) Are Business Cycle Turning Points Uniformly Distributed Throughout the Year? Discussion paper 3891, 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 C.A. Sims (ed.), *Advances in Econometrics—Sixth World Congress*, pp. 257–316. Cambridge, UK: Cambridge University Press.

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., R. McCulloch & R. Tsay (1998) Bayesian inference for a general class of periodic Markov regime-switching models. *Journal of Applied Econometrics*.

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

Hamilton, J.D. (1988) Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control* 12, 385–423.

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.

Hamilton, J.D. (1993) Estimation, inference and forecasting time series subject to changes in regime. In C.R. Rao & G.S. Maddala (eds.), *Handbook of Statistics*, Vol. 11, pp. 231–260. Amsterdam: North-Holland.

Hamilton, J.D. (1994) State-space models. In R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics*, Vol. 4, pp. 3038–3080. Amsterdam: North-Holland.

Hamilton, J.D. (1996) Specification testing in Markov switching time series models. *Journal of Econometrics* 70, 99–126.

Hansen, B.E. (1992) The likelihood test under non-standard conditions: testing the Markov trend model of GNP. *Journal of Applied Econometrics* S61–S82; erratum, 11, 195–198.

Hansen, L.P. (1982) Large sample properties of generalized methods of moments estimators. *Econometrica* 50, 1029–1054.

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

Hansen, L.P. & T.J. Sargent (1996) Recursive Linear Models of Dynamic Economies. Manuscript, Hoover Institution, Stanford University.

Hylleberg, S. (1986) *Seasonality in Regression*. New York: Academic Press.

Karlsen, H.A. (1990) *A Class of Nonlinear Time Series Models*, Ch. 5. Ph.D. Dissertation, University of Bergen.

Kim, C.J. (1994) Dynamic linear models with Markov switching. *Journal of Econometrics* 60, 1–22.

McCulloch, R.E. & R.S. Tsay (1994) Bayesian analysis of autoregressive time series via the Gibbs sampler. *Journal of Time Series Analysis* 15, 235–250.

Osborn, D.R. (1988) Seasonality and habit persistence in a life cycle model of consumption. *Journal of Applied Econometrics* 3, 255–266.

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

Osborn, D.R. & J.P. Smith (1989) The performance of periodic autoregressive models in forecasting seasonal U.K. consumption. *Journal of Business and Economic Statistics* 7, 117–128.

Phillips, K.L. (1991) A two-country model of stochastic output with changes in regime. *Journal of International Economics* 31, 121–142.

Shiller, R.J. (1994) Stock market crashes. *McGraw-Hill Encyclopedia of Economics*. New York: McGraw-Hill.

Sichel, D.E. (1994) Inventories and the three phases of the business cycle. *Journal of Business and Economic Statistics* 12, 269–278.

Tiao, G.C. & M.R. Grupe (1980) Hidden periodic autoregressive-moving average models in time series data. *Biometrika* 67, 365–373.

Tjøstheim, D. (1986) Some doubly stochastic time series models. *Journal of Time Series Analysis* 7, 51–73.

Tjøstheim, D. (1990) Non-linear time series and Markov chains. *Advances in Applied Probability* 22, 587–611.

Todd, R. (1983) *A Dynamic Equilibrium Model of Seasonal and Cyclical Fluctuations in the Corn-Soybean-Hog Sector*. Ph.D. Dissertation, University of Minnesota.

Todd, R. (1990) Periodic linear-quadratic methods for modeling seasonality. *Journal of Economic Dynamics and Control* 14, 763–795.