Discussion Paper 84

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
Research Department
P.O. Box 291
Minneapolis, Minnesota 55480-0291

May 1993

A Time Series Model With Periodic Stochastic Regime Switching

Eric Ghysels

Université de Montréal

*We would like to acknowledge the financial support of the SSHRC and NSERC of Canada as well as the Fonds FCAR of Québec. Part of this work was completed while the author was Visiting Research Fellow at Yale University's Cowles Foundation and Research Fellow at the Institute of Empirical Macroeconomics. 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 would particularly like to thank Jim Hamilton for supplying his data and programs and Rob McCulloch for helping with his Gibbs sampler program. We also benefitted from comments made by many seminar participants. Of course, the author bears full responsibility for any errors. Any opinions, findings, conclusions, or recommendations expressed herein are those of the author and not necessarily those of the National Science Foundation, the University of Minnesota, the Federal Reserve Bank of Minneapolis, or the Federal Reserve System.

ABSTRACT

A general class of Markov switching regime time series models is presented that allows one to estimate the nontrivial interdependencies between different types of cycles which make the economy grow at an unsteady rate. The paper further explores results obtained in Ghysels (1991b) suggesting that the economy transits from recessions to expansions with an uneven propensity throughout the year. It is also built on the work of Hamilton (1989) who proposed a stochastic switching-regime model for GNP and has important connections with hidden periodic structures discussed by Tiao and Grupe (1980) or Hansen and Sargent (1990), for instance. The time series models we present 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 stage of the business cycle. While the model is intrinsically nonlinear and stochastic, it produces a linear representation with seasonal effects that appear to be deterministic. The paper provides an elaborate discussion of the regularity conditions for a well-defined covariance structure including explicit formula for characterizing first and second moments. Finally, we present empirical evidence using U.S. GNP data series which tends to support a periodic structure for switching probabilities. The most significant result is the following : it is found that the seasonal in GNP growth significantly affects switching probabilities for regime switches in the nonseasonal growth of GNP. We also analyze the out-of-sample forecast performance of the different models and find that the models exploiting seasonality in transition probabilities perform best.

Key words : periodic structures, nonlinear time series, seasonality, cyclical movements.

1. INTRODUCTION

It may be convenient to assume that the sources of growth cycles and seasonal fluctuations are independent. In reality, all sources of fluctuations are most likely not independent and, even if they were, economic propagation mechanisms would make it hard to disentangle them.[1] Yet, assuming the independence of cyclical and seasonal fluctuations is often made for the sake of convenience. As we like to focus on cyclical fluctuations, this may be a reasonable abstraction of reality, both for the formulation and/or estimation of structural models as well as the prediction and analysis of cyclical and long-run phenomena with (nonstructural) time series models.

Assuming independence of cyclical and seasonal fluctuations or ignoring, for instance, seasonality even if there is dependence might be convenient. But what if the economy recovers from a slump much easier when seasonals are at their peak? What if bankruptcies tend to be postponed until after a shopping season and hence delay economic activity of going bust? What if credit crunches, financial panics, stock market crashes, etc. tend to cluster around the fall or any other season? Is there any reason to believe this is true? If so, is there any way of modeling such interdependencies in a simple fashion?

The purpose of this paper is to present a class of models that would allow one to estimate parametrically the nontrivial (nonlinear) interdependencies between the different patterns of growth fluctuations just described and to allow one to test for such interdependencies. Hence, this paper breaks with a long established tradition of viewing seasonal fluctuations as separate and orthogonal to all other movements in the economy. Two types of models are considered. The first one we analyze generalizes the periodic switching-regime model described in Ghysels (1991b). It has important connections and also builds further on the switching-regime model of Hamilton (1989) as well as the hidden periodic structures presented by Tiao and Grupe (1980) and Hansen and Sargent (1990). The second class of models we consider is inspired by Filardo (1991) and Diebold et al. (1992) who considered switching-regime models with (stochastic) time-varying transition probabilities using logistic functions.

---

1 This is why in models based on the impulse propagation framework introduced by Slutsky (1927) and Frisch (1933) it is difficult to maintain the univariate orthogonal unobserved component statistical decomposition of time series underlying seasonal adjustment procedures [see Ghysels (1990a) for further discussion].

We provide an elaborate discussion of the stochastic process theory of periodic Markov switching-regime models, spelling out regularity conditions for the existence of a well-defined covariance structure with formula that allow one to characterize first and second moments of any given periodic switching-regime model with time-varying coefficients. We also discuss a simple classical LM test for periodicity which is easy to use and, among other attractive features, has the advantage only to require nonperiodic switching-regime model parameter estimates.

This paper also further explores results obtained by Ghysels (1991b, 1992), where it was argued that the NBER business cycle chronology appeared to exhibit an unequal distribution of cyclical turning points and cycle durations. The empirical section of the paper covers MLE and Bayesian estimation of periodic and time-varying transition probability Markov switching-regime models using US post-WWII GNP data. The empirical results can be summarized as follows : if we allow the transition probabilites to differ every quarter, we obtain boundary parameter estimates. Indeed, large samples would be required to estimate a fully unconstrained model. We consider, in turn, two ways of constraining the transition probabilities throughout the year. One strategy consists of relying on Bayesian techniques to smooth transition probabilities with a common prior for each quarter. While there is enough evidence to suggest heterogeneity of stochastic switching throughout the year, it depends unfortunately quite critically on the formulation of the prior. A second route that is pursued appears more promising and also yields the most significant empirical results of the paper. Namely, we consider a time-varying transition probability structure where the probabilities of switching in GNP are allowed to depend on the seasonal fluctuations in GNP. It is found that the seasonal in GNP growth significantly affects the switching probabilities (at the conventional 5 % level). There is evidence suggesting that during expansions, the probability of switching is inversely related to the season. Hence, in high season it is less likely to enter a recession. In contrast, as recessions do not cover many data points, we could not uncover a statistically significant relation.

The structure of the paper is as follows. Section 2 deals with the stochastic process theory. Estimation and hypothesis testing are covered in section 3 and, finally, empirical results are discussed in section 4. Section 5 deals with seasonal adjustment and forecasting. Section 6 concludes.

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

A general class of periodic Markov switching-regime models is presented in this section. Special cases of this class include the (aperiodic) switching-regime models considered by Hamilton (1988, 1989, 1990), Garcia and Perron (1989), Phillips (1991), McCulloch and Tsay (1992), Albert and Chib (1993), among others, as well as the periodic Markovian switching-regime structure presented by Ghysels (1991b, 1992) which was used to investigate the nonuniformity of the distribution of the NBER business cycle turning points. The discussion will focus first on a simplified illustrative example to present 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 some of the notations as well as the specific model which is an AR(1) stochastic switching-regime model with a periodic Markov chain. Section 2.2 elaborates on the linear ARMA and linear periodic representation of the (nonlinear) stochastic switching regime 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 switching-regime models is presented in section 2.4.

2.1 A univariate AR(1) model as illustrative example

The purpose of this section is to provide motivation and insights by first using a simple model. Namely, consider a univariate time series process denoted $\{y_t\}$. It will typically represent a growth rate of, say, GNP. Moreover, for the moment, it will be assumed the series was seasonally adjusted via a procedure like the X-11 program of the U.S. Bureau of the Census. Furthermore, let $\{y_t\}$ be generated by the following stochastic structure :

$$(2.1) \quad \left[y_t - \mu[(i_t, \mathcal{s}_t)]\right] = \phi\left[y_{t-1} - \mu[(i_{t-1}, \mathcal{s}_{t-1})]\right] + \varepsilon_t$$

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 (2.1) would simply represent a standard linear stationary Gaussian AR(1) model. Instead, we assume that the intercept changes according to a Markovian switching-regime model, following the work of Hamilton (1989). The "state-of-the-world" process is different, however, from that originally considered by Hamilton. In (2.1) we have $s_t \equiv (i_t, \mathscr{s}_t)$, namely the state of the world is described by a stochastic switching regime process $\{i_t\}$ and a seasonal indicator process, i.e., $\mathscr{s}_t = t \bmod(\mathscr{S})$ where $\mathscr{S}$ is the frequency of sampling throughout the year, e.g., $\mathscr{S} = 4$ for quarterly sampling. The $\{i_t\}$ and $\{\mathscr{s}_t\}$ processes interact in the following way, assuming that $i_t \in \{0, 1\}\ \forall\, t$:[2]

(2.2)		0	1
	0	$q(\mathscr{s}_t)$	$1 - q(\mathscr{s}_t)$
	1	$1 - p(\mathscr{s}_t)$	$p(\mathscr{s}_t)$

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

(2.3) $p(\cdot) = \bar{p}$ and $q(\cdot) = \bar{q}$,

then we obtain the standard homogeneous Markov chain model considered by Hamilton. However, if for at least some $\mathscr{s}_t$ 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 $i_t \in \{0, 1\}$, we have a two-state Markovian chain with periodic variation in the transition probabilities, and define the mean shift function :

(2.4) $\mu(s_t) \equiv \mu[(i_t, \mathscr{s}_t)] = \alpha_0 + \alpha_1 i_t\, ,\ \alpha_1 > 0$.

Hence, the process $\{y_t\}$ takes on a mean shift α_0 in state 1 corresponding to $i_t = 0$ and $\alpha_0 + \alpha_1$ in state 2.[3] Equations (2.1) through (2.4) are a version of Hamilton's model with a periodic stochastic switching process. If state 1 is called a recession, as it has a low mean drift and state 2 an expansion, then according to (2.2) we stay in a recession or move to an expansion with a probability scheme that depends on the season.

---

2 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.

3 Note in (2.4) that the mean shift function only depends on i_t and not on $\mathscr{s}$, though in later developments this restriction will be removed.

2.2 Linear ARMA and periodic ARMA representations of a period Markov switching regime 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 switching-regime structure, as represented by (2.1) through (2.4), and the more conventional linear ARMA processes as well as the periodic ARMA models discussed in Tiao and Grupe (1980), Todd (1983, 1990), Osborn (1988), Osborn and Smith (1989) and Hansen and Sargent (1990), among others. Let us perhaps 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 previous section 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 (1990), 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 described in the previous section also has nonlinearly predictable innovations and features of hidden periodicities combined.

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

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

where $\lambda(\mathscr{s}_t) \equiv -1 + p(\mathscr{s}_t) + q(\mathscr{s}_t)$ and hence $\lambda(\cdot) \in \{\lambda^1, ..., \lambda^{\mathscr{S}}\}$ with $\lambda(\mathscr{s}_t) = \lambda^{\mathscr{s}}$ for $\mathscr{s}_t = \mathscr{s}$.

Moreover, conditional on $i_{t-1} = 1$,

(2.6a) $$v_t(\mathscr{s}_t) = \begin{cases} (1 - p(\mathscr{s}_t)) & \text{with probability } p(\mathscr{s}_t) \\ -\, p(\mathscr{s}_t) & \text{with probability } 1 - p(\mathscr{s}_t) \end{cases}$$

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

$$\text{(2.6b)} \quad v_t(\mathscr{s}_t) = \begin{cases} -(1 - q(\mathscr{s}_t)) & \text{with probability } q(\mathscr{s}_t) \\ q(\mathscr{s}_t) & \text{with probability } 1 - q(\mathscr{s}_t) \end{cases}$$

Equation (2.5) 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 $\{i_t\}$ but also $\{\mathscr{s}_t\}$. If all superscripts were absent from equations (2.5) and (2.6), then we would recover the nonperiodic AR(1) representation of Hamilton's stochastic switching-regime model as it appears in equations (2.3) and (2.4) of his 198-9 paper. Equation (2.5) resembles the periodic ARMA models which were discussed by Tiao and Grupe, Todd, Osborn and Hansen, and Sargent, among others, yet, it is fundamentally different in many respects. The most obvious difference, of course, 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 (2.5) and the more standard periodic linear ARMA models have in common. Following Gladysev (1961), we can consider time invariant representations of (2.5) which are built on stacked, skip-sampled vectors of observations. In particular, let us assume we dispose of a sample of length $\mathscr{S}T$, i.e., T number of years. Let us define the stacked vector of seasons which is sampled at an annual frequency :

$$\text{(2.7a)} \quad \underline{i}_\tau \equiv (i_{\mathscr{S}\tau-\mathscr{S}+1}, i_{\mathscr{S}\tau-\mathscr{S}+2}, \ldots, i_{\mathscr{S}\tau})' \qquad \tau = 1, \ldots, T$$

$$\text{(2.7b)} \quad \underline{v}_\tau \equiv (v^1_{\mathscr{S}\tau-\mathscr{S}+1}, v^2_{\mathscr{S}\tau-\mathscr{S}+2}, \ldots, v^{\mathscr{S}}_{\mathscr{S}\tau})'$$

so that τ represents annual time accounting. Following equation (2.5), we can write the DGP for the vector defined in (2.7a) as follows :

$$\text{(2.8)} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\lambda^2 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & \ddots & \ddots & & & \vdots \\ \vdots & & & & & \ddots & \\ 0 & & & & & 1 & 0 \\ 0 & \cdots & & & \cdots & -\lambda^{\mathscr{S}} & 1 \end{bmatrix} \underline{i}_\tau = \begin{bmatrix} 1 - q^1 \\ 1 - q^2 \\ \vdots \\ \\ \vdots \\ 1 - q^{\mathscr{S}} \end{bmatrix} + \begin{bmatrix} 0 & \cdots & \cdots & \lambda^1 \\ \vdots & & & \vdots \\ \\ \\ \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \underline{i}_{\tau-1} + v_\tau .$$

There are two features about (2.8) which we would like to highlight and digress further on. The first is the appearance of seasonal mean shifts, i.e., what is typically called "deterministic seasonality", the second is the basis of a time-invariant Wold MA representation for the (scalar) $\{i_t\}$ process described by (2.5). We shall focus first on the latter, followed by 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 (2.5) and (2.6). It is easy to see that the vector process $\{\underline{i}_\tau\}$ has a covariance stationary representation now (again under suitable regularity conditions discussed later) as the coefficient matrices in (2.8) are time invariant, yet, as we know, of course, each component of the vector still has a different unconditional mean, variance, etc. Through (2.8) one can derive the Wold representation of $\{i_t\}$. This is usually referred to as the Tiao-Grupe formula. As we will be explicitly using this formula, we will briefly discuss it.[4] Assume that the Wold decomposition representation for the vector process $\{\underline{i}_\tau\}$ can be written as follows :

(2.9) $\underline{i}_\tau = \underline{M}\,(L)\,\underline{\omega}_\tau + \underline{\mu}$

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

(2.10) $S_{\underline{i}}(z) = \underline{M}(z)\,\underline{M}(z^{-1})'$.

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

---

4 For a more elaborate discussion, see Tiao and Grupe's original paper of Hansen and Sargent (1990, Chap. 10). Loosely speaking one way of viewing the $\{\underline{i}_\tau\}$ process is to say that it is a discrete state Markov chain with $2^{\mathcal{S}}$ possible states. When $o < q^{s}$, $p^{s} < 1$ for all $s = 1, ..\ ..., \mathcal{S}$, we know such process has a well-defined and unique steady state distribution [see, e.g., Billingsley (1961) or Doob (1953)]. Since Markov chains which have an asymptotically stable distribution are covariance stationary processes, we know from the spectral representation theorem that the $\{\underline{i}_\tau\}$ must have a well-defined spectrum. The $\{i_t\}$ process has an asymptotically stable distribution as well, since we can marginalize out the seasons from the steady state distribution of the $\{\underline{i}_\tau\}$ process. In fact, by using the spectral representation of both $\{i_t\}$ and $\{\underline{i}_\tau\}$, we can easily find the correspondence between the two linear time series representations of the switching regime process, namely one that ignores the hidden periodicity and one that builds on the periodic structure via the vector representation appearing in (2.8).

(2.11) $s_i(z) = Q(z)\, S_{\underline{i}}(z^{\mathscr{S}})\, Q(z^{-1})'$

where $Q(z) = \mathscr{S}^{-1/2}[1 \;\; z \;\; \ldots \;\; z^{\mathscr{S}-1}]$. From (2.11), one can calculate a spectral representation of the $\{i_t\}$ process or we can derive the linear time domain representation of the process.

Now that we have obtained a linear univariate time invariant representation for the stochastic switching regime process i_t, in this case a finite order ARMA process as we shall explain shortly, what can be said about it? First, 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).[5] The periodic nature of autoregressive coefficients pushes the seasonality into annual lags of the AR polynomial emerging from (2.11) and substantially complicates the MA component.

Ultimately, we are, of course, interested in the time series properties of $\{y_t\}$ as it is generated by (2.1) through (2.4) and how its properties relate to linear ARMA and periodic ARMA representations of the same process. Since,

(2.12) $y_t = \alpha_0 + \alpha_1 i_t + (1 - \phi L)^{-1} \varepsilon_t$,

and ε_t was assumed Gaussian and independent, we can simply view $\{y_t\}$ as the sum of two independent processes : namely, a nonlinear time series process $\{i_t\}$ and a linear process $(1 - \phi L)^{-1} \varepsilon_t$. Clearly, all 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 has the following linear time series representation :

(2.13) $s_y(z) = \alpha_1^2\, s_i(z) + \left|\frac{1}{1 - \phi z}\right|^2 \sigma^2/2\pi$.

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 in Ghysels (1991a).

This linear representation has hidden periodic properties which can be derived from (2.10) 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 (2.9), or alternatively to (2.8). We observe that the linear representation has seasonal mean shifts that would apper as a "deterministic seasonal" in the univariate representation of y_t. Hence, besides the spectral density properties appearing in (2.13), which may or may not show peaks at the seasonal frequency, we note that what looks like seasonal dummies appear in the univariate representation. This result is, of course, quite interesting since 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. While it is a purely random and nonlinear model, the periodicity in the Markov chain makes linear representations appear to have a deterministic seasonal mean shift.

2.3 Some properties of interest

So far, we established some of the characteristics of the stochastic switching regime AR(1) process with periodic transition probabilities. In particular, in the previous 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 some of the stochastic properties of the processes that are of interest. Three properties will be of special interest and thus highlighted in this section. They are : (1) seasonal conditional asymmetries; (2) the periodic duration distribution; and (3) the seasonal impulse response functions. We shall discuss each of these separately.

(1) Seasonal conditional asymmetries

Consider the conditional variance of the innovation process appearing in (2.5). It can be written as :

$$(2.14)\quad E[(v_t^{\delta})^2 \mid i_{t-1}, \delta_t] = \begin{cases} p(\delta_t)\,(1 - p(\delta_t)) & \text{if } i_{t-1} = 1 \\ q(\delta_t)\,(1 - q(\delta_t)) & \text{if } i_{t-1} = 0 \end{cases} .$$

We observed that the variance of the stochastic switching regime process, whether it be presented as a scalar or vector, displays heteroskedasticity, conditional not only with regard to the season but also to the regime shifts. The former source of heteroskedasticity, namely the seasonal variation in (conditional) second moments, is a natural byproduct of the hidden periodicity and also features in the processes studied by Tiao and Grupe, Todd, Osborn, Hansen and Sargent, among others. However, what is different is the asymmetry in conditional second moments blended with the periodic structure.[6] In section 2.4, we will further elaborate on these issues.

The seasonal asymmetry in the first moments mentioned earlier, as advanced and empirically documented in Ghysels (1990b), also emerges in some sense from our model, though only in the following conditional way :

$$(2.15)\ E(y_t \mid i_{t-1}, \mathscr{s}_t) = \begin{cases} \alpha_0 + \alpha_1\, p(\mathscr{s}_t) & \text{if } i_{t-1} = 1 \\ \alpha_0 + \alpha_1(1 - q(\mathscr{s}_t)) & \text{if } i_{t-1} = 0 \end{cases}$$

so that the size of the (predicted) seasonal mean will depend on whether we are in either one of the two states. In section 2.4, we shall introduce other forms of asymmetries in seasonal first moments.

(2) Periodic duration distribution

This feature highlights a characteristic proper to periodic Markov chains that was exploited in Ghysels (1990b) to test the presence of periodicity via exact small sample rank-based nonparametric statistics. When a Markov chain is periodic, as in (2.5), then the distribution of the length of time spent in any particular regime depends on the starting season.

Depending on whether the switch into a new regime occurred in the first, second or any other season will affect the expected length of the regime's lifetime. To see this, conditional on being in, say, state $i_t = 0$, the expected duration of the low growth state will be :

---

6 The notion of "seasonal asymmetries", i.e., seasonal patterns which do not appear the same during expansions and recessions was advanced in Ghysels (1990c). While the emphasis was slightly different in that paper as it dealt with unconditional asymmetries and was almost exclusively on seasonal mean shifts, i.e., first moments, the possibility of asymmetries in other moments was also raised. The AR(1) periodic switching regime indeed yields asymmetries in the seasonal conditional heteroskedasticities.

$$\sum_{k=1}^{\infty} k \left[\prod_{j=1}^{k-1} \left[\sum_{\mathscr{s}=1}^{\mathscr{S}} (1_{\mathscr{s}t+j}) \, q(\mathscr{s}_t) \right] \right] \left[\sum_{\mathscr{s}=1}^{\mathscr{S}} (1_{\mathscr{s}t+k}) \, (1 - q(\mathscr{s}_t)) \right] .$$

For the purpose of illustration, take the simple example where $\mathscr{S} = 2$, i.e., when only two seasons occur, then the expected durations are respectively :

$$\sum_{k=1}^{\infty} (q(1)\, q(2))^{k-1} \left[(1 + 2k)\,(1 - q(1)) + (2 + 2k)\,(1 - q(2))\, q(1)\right]$$

$$\sum_{k=1}^{\infty} (q(2) q(1))^{k-1} \left[(1 + 2k)\,(1 - q(2)) + (2 + 2k)\,(1 - q(1))\, q(2)\right]$$

which will differ for $q(1) \neq q(2)$. The dependence of the duration distribution on starting seasons using the NBER Business cycle chronology was studied in detail in Ghysels (1990b).

(3) Seasonal impulse response functions

The purpose here is only to draw attention to the fact that due to the hidden periodicity, there is, in fact, also a periodic impulse response scheme that goes with the Wold decompositions conditional on the season as presented in (2.11). Hansen and Sargent (1990) study in detail how the impulse response mechanisms operate in a periodic (linear) environment. We can only refer the reader here to the detailed exposition they presented. It should also be emphasized that Hansen and Sargent provide several examples of economic structural models which yield a linear periodic representation. Similar attempts were made by Todd (1983, 1990) and Osborn (1988), though Hansen and Sargent provide a unifying general equilibrium approach.

2.4 A general class of periodic Markov stochastic switching-regime models

Having mostly relied on intuition and a specific example so far, we will now turn our attention to generalizations. All technical material is conveniently presented at the end of the paper in Appendix A.1. Here, we shall only point to the different directions in which one can generalize the model and discuss how they can be formally treated.

Consider the set $\mathscr{Y}$ of $\mathbb{R}^n$-valued discrete time vector processes defined on the probablity space $(\Omega, \mathscr{A}, \mathscr{P})$ where for each $\omega \in \Omega$, $\{y_t(\omega)\} \in \mathscr{Y}$ is generated as follows :[7]

$$(2.16)\ y_t = b_0(i_t, s_t, z_t)\, x_{0t} + \sum_{j=1}^{\ell} b_j(i_t, s_t, z_t) \left[y_{t-j} - b_0(i_{t-j}, s_{t-j}, z_{t-j})\, x_{0t-j}\right] + \delta_t\,.$$

Equation (2.16) 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, s_t, z_t), where $\{i_t\}$ follows a Markov chain with transition probability matrix $P(s_t, z_t)$ and $s_t \equiv t \bmod \mathscr{S}$ as defined before, while z_t is a set of variables affecting the transition probabilities similar to Filardo (1991) and Diebold et al. (1992). The regressors x_{0t} in equation (2.16) are fixed, namely consisting of either a constant or a constant and $\mathscr{S} - 1$ seasonal dummies, while the error process δ_t is i.i.d. $N(0, \Lambda)$.

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 each of r possible regimes. As there are ℓ lags in equation (2.16), the Markov process will have $r^{\ell+1}$ states *each season.* Hence, the Markov chain throughout the year is described by the set $\{P(s, z_t),\ s = 1, \ldots \mathscr{S}\}$ where $P(s, z_t)$ is a $r^{\ell+1} \times r^{\ell+1}$ transition probability matrix. Following Filardo (1991) and Diebold et al. (1992), one can consider the transition probabilities to be time-varying, evolving as logistic functions of $z_t' \gamma_i(s)$, $s = 1, \ldots, \mathscr{S}$. Hence, in state i, a different vector $\gamma(\cdot)$ applies to each season. To illustrate this further, just let $\ell = 0$ and $r = 2$ for the moment. Then we have :

$$(2.17)\quad \begin{array}{ccc} & 0 & 1 \\ 0 & \dfrac{\exp(z_t'\gamma_0(s_t))}{1 + \exp(z_t'\gamma_0(s_t))} & 1 - \dfrac{\exp(z_t'\gamma_0(s_t))}{1 + \exp(z_t'\gamma_0(s_t))} \\ 1 & 1 - \dfrac{\exp(z_t'\gamma_1(s_t))}{1 + \exp(z_t'\gamma_1(s_t))} & \dfrac{\exp(z_t'\gamma_1(s_t))}{1 + \exp(z_t'\gamma_1(s_t))} \end{array}.$$

---

7 To save on notation again, we substitute y_t for $y_t(\omega)$, etc.

A special case of (2.17) is where z_t is just a constant. Then the transition matrix simply becomes a function of s_t only, which is what appeared in (2.2). As $\gamma_i(\cdot)$ becomes a scalar in such a case, one can simply express the transition probabilities $p(\cdot)$ and $q(\cdot)$ in (2.2) via the logistic function $q(s_t) = \exp(\gamma_0(s)) / (1 + \exp(\gamma_0(s))$ and $p(s) = \exp(\gamma_1(s)) / (1 + \exp(\gamma_s(s))$. Another special case of (2.17) which is also of particular interest for empirical work is the case where $P(s_t, z_t) \equiv P(z_t)$, i.e., independent of s_t. Of course, this corresponds exactly to the analysis in Diebold et al. (1992). Yet, unlike Filardo or Diebold et al., it is important to observe that z_t may contain seasonal processes. Hence, the transition probability matrix becomes *stochastic* and seasonal instead of *deterministic* and seasonal, as in the case in the purely periodic transition scheme of section 2.1.

The matrix functions $b_1(\cdot)$ through $b_\ell(\cdot)$ appearing in (2.16) are $n \times n$ polynomial lag matrices which 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}, i.e., $x_{0t} = 1 \ \forall\ t$, and $b_0(i_t, s_t)$ only depending on $\{i_t\}$, we recover the most familiar case where $\{y_t\}$ is driven by a stochastic mean shift which is a function of a latent Markov process determining the regime switches. It should be noted though that y_t can be a $n \times 1$ vector process, so that $b_0(\cdot)$ determines a $n \times 1$ vector of mean shifts depending on $\{i_t\}$ for the joint multivariate process. When x_{0t} also includes seasonal dummies, several new issues emerge. In particular, until now, we have not really made explicit whether the $\{y_t\}$ process is seasonally adjusted or not, using conventional methods such as the X-11 procedure of the U.S. Bureau of the Census. The effect of such filters on a process generated by (2.16) is a matter of discussion which will be deferred to later. For the moment, suffice it to say that such filters are usually inspired on and defined in terms of linear time series unobserved components which do not involve any periodic or nonlinear features. To bypass these questions, let $\{y_t\}$ be an unadjusted series. We noted in section 2.2 that the periodic nature of the Markov chain already induced a "seasonal dummy"-type behavior in a linear representation where the seasonal mean shifts are entirely governed by the transition probabilities $P(s)$, $\forall s$. Yet, as these transition probabilities determine quite a lot of features ranging from seasonal asymmetries in conditional variance to periodicity in duration length, as discussed in section 2.3, it is reasonable to introduce seasonal dummies separately. It not only relaxes the burden of a tightly parameterized model, but more importantly leads to testable hypotheses.

Augmenting the model with seasonal dummies, makes $b_0(\cdot)$ a $n \times \mathscr{S}$ matrix function. An interesting proposition, already mentioned in section 2.3, is whether there are seasonal asymmetries in mean shifts, that is to say, coefficients in $b_0(\cdot)$ corresponding to the seasonal dummies which depend on $\{i_t\}$. If so, then seasonal mean shifts depend on the stage of the business cycle, as suggested in Ghysels (1990c).[8]

Obviously, equation (2.16) contains many features all at once, making it potentially a richly parameterized model that will be too demanding for most data sets to be inferred from. As $b_1, ..., b_\ell(\cdot)$ and $b_0(\cdot)$ are allowed to depend on $\{i_t\}$ and $\{\mathscr{s}_t\}$, one can indeed produce some quite complex dynamics in polynomial lags, seasonals and regimes.[9]

So far, we have presented a vector stochastic switching-regime process with possibly seasonal transition properties, both periodic through $\mathscr{s}_t$ 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? Discussing regularity conditions requires a fair amount of notation and elaboration. To avoid interrupting the flow of the paper, we refer the reader to Appendix A.1 which contains a formal discussion of regularity conditions under which the processes will be well behaved, that is to say, asymptotically stable and having at least finite second moments. It should be noted though that our formal treatment only covers the case where z_t is a constant, hence the transition matrix $P(\cdot)$ is nonrandom, yet possibly periodic.[10] It is shown that (periodic) Markov switching-regime processes can be treated as doubly stochastic vector AR(1) processes, using the terminology coined by Tjøstheim (1986). Our formalization then

---

8 One parameterization, leading to a convenient testable hypothesis would be as follows : $b_0(i_t, \mathscr{s}_t) \equiv (\alpha_0 + \alpha_1(i_t), \alpha_{01} + \alpha_{11}(i_t), ..., \alpha_{0\,\mathscr{S}-1} + \alpha_{1\,\mathscr{S}-1}(i_t))$ where $i_t \in \{0, 1\}$ $\forall t$ and $\alpha_{1j}(i_t) = \alpha_{1j} i_t$. Hence, for $\alpha_{1j} = 0$ $\forall j = 1, ... \mathscr{S} - 1$, none of the seasonal mean shifts depend on $\{i_t\}$, e.g., the stage of the business cycle. In contrast, when for some j $\alpha_{1j} \neq 0$ then depending on whether $i_t = 0$ or 1, the seasonal mean shift will be $\alpha_{0j} + \alpha_{1j}$ respectively.

9 See, for instance, Hansen (1991) and McCulloch and Tsay (1992) for switching regime models with state-dependent AR polynomials.

10 A formal treatment of such regularity conditions has been absent from the literature on Hamilton-type models. Since periodic Markov switching-regime models cover as special cases an aperiodic homogeneous Markov scheme, it should be noted that the regularity conditions apply to a large set of applications hitherto treated informally.

relies on Tjøstheim (1990) and Karlsen (1990) to characterize the necessary conditions for weak stationarity. A general linear representation is discussed highlighting attributes such as the appearance of seasonal mean shifts and periodic autocovariance structure.

3. ESTIMATION AND HYPOTHESIS TESTING

Estimating Markov switching-regime model is mostly likelihood based, either via classical methods like in Hamilton (1989) among others, or via Bayesian methods following the work of Albert and Chib (1991) and McCulloch and Tsay (1992). We will proceed along those same lines, namely focus on classical MLE as well as Bayes estimation using a Gibbs sampler approach.[11] This section is structured as follows. We shall first discuss the formulation of the likelihood function in section 3.1. The next section is devoted to classical hypothesis testing. Anticipating some of the empirical results, we shall devote special attention to "boundary problems" in section 3.2 as well. The Bayesian approach to estimation is covered in section 3.3.

3.1 The formulation of the likelihood function

The estimation of Markov switching-regime models is covered in detail in Hamilton (1989, 1991b) for the case where the Markov chain is homogeneous and in Diebold et al. (1992) for time-varying transitions. Therefore, we shall restrict ourselves here to a brief discussion, only highlighting new features occurring because of periodicity. In general, we seek to estimate the parameter vector θ governing the coefficient matrices $b_j(\cdot)$, $j = 0, 1, \ldots, \ell$ and the covariance matrix Λ from equation (2.16). We will make some concessions regarding generality and focus instead on the special case of two primitive states with a simple periodic Markov chain, like in (2.2), involving a scalar stochastic process y_t, i.e., $n = 1$. Given a sample of size $\mathscr{S}T$, i.e., T full years of data points, the log-likelihood function can be written as :

$$(3.1)\quad L(Y_{\mathscr{S}T}, \theta) \equiv \sum_{t=1}^{\mathscr{S}T} \log p(y_t \mid Y_{t-1}; \theta)$$

11 It should be noted that a method-of-moments approach based on the results obtained in Appendix A.1 (Theorem A.1) is also a feasible estimation strategy. A full discussion of such an approach is beyond the scope of this paper.

where $p(\cdot \mid Y_{t-1}; \theta)$ represents the probability distribution of y_t, given observations up to t − 1, i.e., $Y_{t-1} \equiv \{(y_{t-j}, \delta_{t-j}), j \geq 1\}$. Hamilton (1989, 1991b) goes into detail how to formulate $p(\cdot \mid Y_{t-1}; \theta)$ 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 $s_t = (\xi_t, \delta_t)$ at any given point in time. This probability can be written as :

$$(3.2) \quad p(s_t = (i_t, \delta_t) \mid Y_j; \theta)$$

where j can be smaller, equal or greater than t. The algorithm starts out with the unconditional probability. The first observation is drawn from π_δ with $\delta = \delta_1$ and π_δ being determined by (2.31). When $i_t \in \{0, 1\}$, the initial unconditional probability would be, using (2.2) to describe the transition probabilities :

$$(3.3) \quad p(s_t = (1, \delta_1)) = \frac{1 - p(\delta_1)}{(1 - q(\delta_1) + 1 - p(\delta_1))} .$$

Furthermore, the joint probability of s_1 and s_2 can be written as

$$(3.4) \quad p(s_2 = (1, \delta_2), s_1 = (1, \delta_1)) = p(\delta_1) \frac{1 - p(\delta_1)}{(1 - q(\delta_1) + 1 - p(\delta_1))} ,$$

Iterations similar to (3.4) can be computed for an initial segment $t = 1, ..., \ell + 1$, where ℓ is the lag length of the AR polynomial in (2.16). Then, the density of the $\ell + 1$ sample points conditional on $(s_1, ..., s_{\ell+1})$ can be written as :

$$(3.5) \quad p(Y_{\ell+1} \mid s_{\ell+1}, ..., s_1; \theta) = (2\pi)^{(\ell+1)/2} |\Lambda^\ell|^{-\frac{1}{2}} \times$$

$$\exp\{(\tilde{y}^\ell_{\ell+1} - B^\ell(\xi_{\ell+1}, \delta_{\ell+1})\, \tilde{y}^\ell_\ell)' (\Lambda^\ell)^{-1} (\tilde{y}^\ell_{\ell+1} - B^\ell(\xi_{\ell+1}, \delta_{\ell+1})\, \tilde{y}^\ell_\ell)\}$$

where $\tilde{y}^\ell_\ell$ is as defined in (A.1.6). Then, the conditional distribution of the $\ell + 1$ first states, given data points in $Y_{\ell+1}$ can be expressed as :

(3.6) $p(s_{\ell+1}, \ldots, s_1 \mid Y_{\ell+1}) =$

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

using (3.4) and (3.5) to formulate the conditional and unconditional densities appearing in (3.6). The latter starts off an algorithm that is applied iteratively throughout the sample, beginning with :

$$(3.7)\quad p(s_{t+1} = (k, \delta_{t+1}), s_t = (j, \delta_t) \mid Y_t) = [\sum_{\delta=1}^{\mathcal{S}} 1(\delta_{t+1} = \delta)\, p_{kj}^{(\delta)}]\, p(s_t \mid Y_t)$$

where $1(\delta_t = \delta)$ is a seasonal indicator function used quite extensively in the remainder of the paper, namely

$$1(\delta_t = \delta) = \begin{cases} 1 & \text{if } \delta_t = \delta \\ 0 & \text{otherwise} \end{cases}$$

Next, one can write :

(3.8) $p(y_{t+1} \mid s_{t+1}, s_t, Y_{t+1}) \sim$

$$N(b_0(i_{t+1}, \delta_{t+1})\, x_{0t+1} + \sum_{j=1}^{\ell} b_j(i_{t+1}, \delta_{t+1})\, [(y_{t-j+1} - b_0(i_{t-j+1}, \delta_{t-j+1})\, x_{0t-j+1}], \Lambda)$$

Finally, yielding

(3.9) $p(s_{t+1} = (k, \delta_{t+1}) \mid Y_{t+1}) =$

$$\sum_{j=1}^{K} \frac{p(y_{t+1} \mid s_{t+1}, s_t, Y_t) \left[\sum_{\delta=1}^{\mathcal{S}} 1(\delta_{t+1} = \delta)\, p_{kj}(\delta)\right] p(s_t = (j, \delta_t) \mid Y_t)}{\left[\sum_{u=1}^{K} \sum_{v=1}^{K} p(y_{t+1} \mid s_{t+1}, s_t, Y_t) \left[\sum_{\delta=1}^{\mathcal{S}} 1(\delta_{t+1} = \delta)\, p_{uv}(\delta)\right] p(s_t = (v, \delta_t) \mid Y_t)\right]}.$$

The expression in (3.9) together with (3.1) yields the desired log-likelihood function. One important feature about (3.9) will be most useful in the next section, in particular with respect to the derivation of an LM statistic. Namely, it should be noted that only the $p(\mathcal{s})$ for $\mathcal{s}_{t+1} = \mathcal{s}$ appears in the recursion formula for $p(s_{t+1} \mid Y_{t+1})$. All other transition matrices $p(\mathcal{s})$ with $\mathcal{s} \neq \mathcal{s}_{t+1}$ do not appear directly, though they affect, of course, $p(s_t \mid Y_{t+1})$ on the right-hand side of (3.9).

3.2 Classical hypothesis testing

Based on the log-likelihood function specified in the preceding section, we can construct a number of tests. As general specification tests were developed elsewhere, we will not devote much attention to their presentation. Indeed, Hamilton (1991c) developed tests for omitted autocorrelation, omitted ARCH and misspecification of the Markovian dynamics. Such tests can easily be applied to the framework of section 2. Instead, we will focus exclusively our attention on the principal hypothesis of interest, namely the periodicity of the Markov structure appearing in (2.2). The hypothesis of no periodic structure can be formally stated as follows :

$$(3.10)\ H_0 : p_{ij}(\mathcal{s}) = \bar{p}_{ij} \qquad \forall\, i, j \in \{1, \dots, K\},\ \mathcal{s} = 1, \dots, \mathcal{S}.$$

This hypothesis is "standard", and hence does not involve issues like testing when nuisance parameters are not identified under the null and issues which emerge when testing model (2.16) against a linear time series model, for instance.[12]

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

$$(3.11)\ LR = -2[L(y_T, \hat{\theta}_c) - L(y_T, \hat{\theta}_u)] \xrightarrow{d} \chi^2\,(df)$$

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 $(\mathcal{S} - 1) \times K$.

---

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. It is beyond the scope of our paper to reassess Hansen's finding assuming Hamilton's model exhibits periodic Markovian features.

Next, we consider a LM test for the same hypotheses. A LM test has several advantages over the LR test. We will only estimate the restricted, i.e., nonperiodic model, and evaluate 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. But 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. In the previous section, it was shown that the conditional probability $p(s_{t+1} \mid Y_{t+1})$ appearing in (3.9) only involves $p_{ij}(s)$ for $s_{t+1} = s$. 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, i.e., K = 2, we have for i = 1, 2 :[13]

(3.12)
$$r_{it}^{s}(\lambda) \equiv \partial \log p(y_t \mid Y_{t-1}; \lambda) / \partial p_{ii}(s) =$$

$$1(s_t = s)[(p_{ii}(s))^{-1}\, p\,(s_t = (i, s_t), s_{t-1}(i, s_{t-1}) \mid Y_t)$$

$$+ (1 - p_{ii}(s))^{-1}\, p(s_t = (j, s_t), s_{t-1} = (i, s_{t-1}) \mid Y_t)]$$

$$+ (p_{ii}(s))^{-1} \left\{ \sum_{n=2}^{t-1} 1(s_n = s)[p(s_n = (i, s_n), s_{n-1} = (i, s_{n-1}) \mid Y_t) \right.$$

$$- p(s_n = (i, s_n), s_{n-1} = (i, s_{n-1}) \mid Y_t)]$$

$$+ (1 - p_{ii}(s))^{-1} \left\{ \sum_{n=2}^{t-1} 1(s_n = s)[p(s_n = (j, s_n), s_{n-1} = (i, s_{n-1}) \mid Y_t) \right.$$

$$\left. - p(s_n = (j, s_n), s_{n-1} = (i, s_{n-1}) \mid Y_t)] \right\}$$

$$+ 1(s_1 = s)\, \frac{p(s_1 = (i, s_1) \mid Y_t) - p(s_1 = (i, s_1) \mid Y_{t-1})}{1 - p_{ii}(s)}$$

for $t \geq 2$ while for $t = 1$:

---

13 The details of the derivations are omitted here : they appear in Ghysels and Hall (1992) for the general nonlinear standard asymptotic distribution theory case. Linear systems with I(1) processes are also dealt with in that paper.

$$(3.13)\ r^{\mathscr{s}}_{i1}(\lambda) = 1(\mathscr{s}_1 = \mathscr{s}) \frac{p(s_1 = (i, \mathscr{s}_1)|Y_1) - [(1 - p_{jj}(\mathscr{s}))/(1 - p_{ii}(\mathscr{s}) + 1 - p_{jj}(\mathscr{s}))]}{1 - p_{ii}(\mathscr{s})}$$

where j = 2 when i = 1 and j = 1 for i = 2. From (3.13) and (3.14), we define :

$$(3.14)\ R^{\mathscr{s}}_t(\lambda) \equiv \begin{bmatrix} r^{\mathscr{s}}_{1t}(\lambda) \\ r^{\mathscr{s}}_{2t}(\lambda) \end{bmatrix}$$

and $R_t(\lambda) \equiv (R^1_t(\lambda)' \ldots R^{\mathscr{S}}_t(\lambda)')'$. The latter is a stacked system of score functions for each season involving only transitions from that season. The score test can then be formulated as :

$$(3.15)\ LM^M_P \equiv \mathscr{S}T[(1/\mathscr{S}T) \sum_{t=1}^{\mathscr{S}T} R_t(\hat{\lambda}_c)]'[(1/T) \sum_{t=1}^{\mathscr{S}T} R_t(\hat{\lambda}_c)\, R_t(\hat{\lambda}_c)']^{-1} [(1/\mathscr{S}T)$$

$$\sum_{t=1}^{\mathscr{S}T} R_t(\hat{\lambda}_c)] \xrightarrow{d} \chi^2((\mathscr{S} - 1) \times K) .$$

3.3 Bayesian inference using the Gibbs sampler

In a series of recent papers Albert and Chib (1991) and McCulloch and Tsay (1992) discuss Bayesian estimation of Hamilton-type model using Gibbs sampling techniques. The usefulness of Gibbs sampling as a simulation tool has been shown in many statistical applications, see, e.g., Geman (1984), Gelfand et al. (1990), Gelfand and Smith (1990) and Casella and George (1992), and econometric applications, see, e.g., Chib (1991), Geweke (1991a, b, 1992), McCulloch and Rossi (1992) among others. The appeal of the Gibbs sampler is twofold, first like Hamilton's algorithm, it also avoids direct calculation of the likelihood function and second, it is particularly suited for computing joint posterior distributions from conditional distributions of subsets of the parameters. The basic idea to make the Gibbs sampler work is to view the unobserved state process $\{\xi_t\}$ as missing data points through the entire sample and treat them alongside the unknown parameters to compute posterior densities.

Our setup follows closely the approach pursued by McCulloch and Tsay (1992), though their formulation is more general and includes Hamilton's model as a special case. Again, with the particular application in mind covered in the next section, we tailor their setup to our specific model specification used for the empirical investigation.[14] All details are deferred to Appendix A.2 except for an important key aspect of the formulation of the prior, namely the prior for the parameters governing the periodic Markov chain. We confine our attention again to a two-state process, i.e., r = 2, then adopting the notation of (2.2), the priors are formulated as follows :

$$(3.16)\ 1 - q(s) \sim \text{Beta}(a_{01}, a_{02}) \qquad s = 1, .., \mathcal{S}$$

$$(3.17)\ 1 - p(s) \sim \text{Beta}(a_{11}, a_{12}) \qquad s = 1, .., \mathcal{S}$$

Hence, we attribute the same prior mean and dispersion throughout the year in each of the two primitive states. This prior formulation seems fairly natural, not favoring any particular season a priori, and extends the standard Bayesian Beta distribution prior to the case of periodic Markov chains.

4. EMPIRICAL RESULTS

We now turn our attention to estimation and testing Markov switching models with periodic or stochastic seasonal transition features using U.S. post-WWII GNP data. Before estimating the (nonlinear) Markov switching model, we study some of its implications as they emerged from the stochastic properties discussed in section 2. A first subsection is devoted to this type of preliminary analysis. Then, we move on to ML estimation of Markov switching-regime models in section 4.2. Section 4.3 covers Bayesian inference using the techniques described in section 3.3. A final section 4.4 is devoted to forecasting, seasonal adjustment and other matters indirectly related to the switching-regime model.

4.1 On periodic features

The purpose of this section is to conduct a preliminary and exploratory analysis. Namely, we test whether the linear representation of GNP series exhibits periodic features. In section 2, it was noted, both with the illustrative example and the general framework, that a periodic stochastic switching-regime model implies a periodic

---

14 Bayesian statistical inference for a general class of periodic Markov switching regression models based on the Gibbs sampler is discussed in detail in Ghysels, McCulloch and Tsay (1992).

structure in the linear representation. Hence, a first task will be to test periodicity in linear representations of GNP series. The second task will be to study seasonal patterns in the nonlinear predictability of the linear projection innovations.

Table 4.1 : Wald Tests of Linear Periodic Structures in U.S. Post–WWII GNP

Seasonally Unadjusted Data 51:1–84:4

$$\text{Estimated model} : \Delta \log GNP_t^{NSA} = \alpha_0 + \sum_{s=2}^{4} 1(s_t = s)\alpha_s$$

$$+ \sum_{j=1}^{8} \theta_j \Delta \log GNP_{t-j}^{NSA} + \sum_{j=1}^{8} \sum_{s=2}^{4} \theta^*_{sj} (1(s_{t-j} = s) \Delta \log GNP_{t-j}^{NSA})$$

	Q2	Q3	Q4
lags 1–8	0.09	0.32	0.00
lags 1–4	0.30	0.36	0.33
lags 5–8	0.22	0.40	0.00

Joint tests of all seasons : lags 1–8 : 0.00 / lags 1–4 : 0.38 / lags 5–8 : 0.00

Tests for periodicity in linear models for U.S. GNP appear in Table 4.1. The results reported were obtained from unadjusted series. For the sake of convenience, we report Wald tests for periodicity, i.e., estimate an AR(p) model with the cross–product of GNP growth series and seasonal dummies included and test zero restrictions on their coefficients.[15] The model specifications appear explicitly in Table 4.1. It should be noted that the model specification with unadjusted data is one where the growth of GNP is projected on a constant and seasonal dummies and lags of GNP that ignore the possibility of roots on the unit circle at the seasonal frequencies.[16] Of course, the linear specification adopted is one that follows from the periodic stochastic switching regime specification presented in section 2.

---

15 For applications of Wald tests for periodicity, see, e.g., Osborn (1988) and Franses (1992).

16 An important issue is whether U.S. GNP series are better described by a model with roots on the unit circle at the seasonal frequencies. It appears that for the sample covered in Table 4.1, there is evidence of a root at the seasonal bi–annual frequency. This result is fairly robust, in the sense that tests proposed by Hylleberg, Engle, Granger and Yoo (1990) applied to GNP series yield the same conclusion, regardless of the AR augmentation chosen or sample sizes typically used [see, for instance, Ghysels, Lee and Siklos (1992) for a more elaborate discussion]. One should note though that the use of the tests proposed by Hylleberg, Engle, Granger and Yoo, are not explicitly constructed to have power against periodic alternatives. See Ghysels and Hall (1992) for further discussion of this point.

Let us first concentrate on the unadjusted series covering the 51:1-84:4 sample. Hence, this sample is the same as that covered by Hamilton (1989), yet unadjusted rather than adjusted series are investigated. There is a fair amount of evidence favoring a periodic structure in the linear representation, as shown in Table 4.1. Joint tests involving lags 1-8 or lags 5-8 reject the null of periodicity, particularly for Q4. The seasonal adjustment procedure destroys the periodic pattern, however. Though not reported in Table 4.1, we also conducted the same tests with the adjusted data used by Hamilton, for instance, and found no evidence for periodicity.[17] Hence, if any periodicity is left after filtering with X-11, it must appear in nonlinear features of the data. Let us therefore investigate such features, without actually estimating the Markov model. Namely, we turn our attention to regression models involving $p_t = (i_t = 1 \mid Y_t)$ computed from Hamilton's aperiodic model using seasonally adjusted data.[18] The following regression is obtained :

$$\Delta \log GNP_t^{SA} = \underset{(0.293)}{-0.194} \; \underset{(0.164)}{-0.070} \, \Delta \log GNP_{t-1}^{SA} \; \underset{(0.101)}{-0.017} \, \Delta \log GNP_{t-2}^{SA}$$

$$\underset{(0.093)}{-0.152} \, \Delta \log GNP_{t-3}^{SA} \; \underset{(0.092)}{-0.195} \, \Delta \log GNP_{t-4}^{SA} \; \underset{(0.643)}{+1.527} \, p_{t-1}$$

$$\underset{(0.304)}{+0.111} \, p_{t-1}^{(2)} \; \underset{(0.300)}{-0.106} \, p_{t-1}^{(3)} \; \underset{(0.304)}{+0.672} \, p_{t-1}^{(4)}$$

$$R^2 = 0.24 \qquad D.W. = 2.03$$

where $p_t^{(i)} = 1(\delta_t = i)\, p_t$. If we ignore the fact that p_t is a generated regressor and just use significance levels from the estimated standard errors, we not only find, as Hamilton did, nonlinear predictability in the sense that p_{t-1} is significant, we also find a significant individual seasonal effect for Q4. Moreover, the joint test that $p_{t-1}^{(i)} = 0$, $i = 2, \dots, 4$ has a p-value of 0.06. Of course, one has to discount the fact that p_t is a generated regressor, as noted before, yet it is remarkable how such a result is obtained with seasonally adjusted data. It is, of course, easy to understand following the

---

17 Asymptotically, one should still find evidence of periodicity after linear filtering, yet the power of tests in sample size commonly used is dramatically reduced because of seasonal adjustment [see Ghysels and Hall (1992) for further discussion].

18 Please note that the state of the world is simply described by i_t as δ_t does not figure in such a model.

discussion in section 2 regarding the stochastic features of a periodic Markov switching-regime model. Hamilton attributed the significant coefficient for p_{t-1} as supporting evidence of the nonlinear Markov switching-regime model. Likewise, one can make a similar argument for the periodic model on the basis of this preliminary investigation.

Hamilton also investigated whether the squared estimated residuals obtained from an AR(4) model for $\Delta \log GNP_t^{SA}$ were predictable using p_{t-1}. Again, such finding would be supportive of a Markov switching structure. Would there be any periodicity in this relationship? Consider the following regressions :

$$\hat{\varepsilon}_t^2 = \underset{(0.299)}{1.564} \; \underset{(0.370)}{-0.805}\, p_{t-1}$$

$$\hat{\varepsilon}_t^2 = \underset{(0.300)}{1.587} \; \underset{(0.441)}{-0.529}\, p_{t-1} \; \underset{(0.431)}{-0.785}\, p_{t-1}^{(2)} \; \underset{(0.427)}{-0.279}\, p_{t-1}^{(3)} \; \underset{(0.441)}{-0.116}\, p_{t-1}^{(4)}$$

where $\hat{\varepsilon}_t^2$ is the estimated residual squared from an AR(4) model for $\Delta \log GNP_t^{SA}$. It is interesting to note that p_{t-1} becomes insignificant once the $p_{t-1}^{(i)}$ are included and that $p_{t-1}^{(2)}$ seems to carry the predictive power for $\hat{\varepsilon}_t^2$ rather than the overall p_{t-1} series. Here, again, the periodic features transpire, despite seasonal adjustment of the series. Obviously, seasonal adjustment filters, while destroying all of the linear productivity do not completely phase out nonlinear periodicity.

4.2 Classical maximum likelihood estimation of Markov switching models

We consider two types of models. Both relate to the general framework discussed in section 2.5. The first model is a simple periodic version of Hamilton's original model, as discussed throughout section 2. The AR polynomial expansion is of order 4, hence :

$$(4.1) \quad \Delta \log GNP_t^{SA} = \alpha_0 + \alpha_1 i_t + \sum_{j=1}^{4} (\Delta \log GNP_{t-j}^{SA} - \alpha_0 - \alpha_1 i_{t-j}) + \varepsilon_t$$

with $i_t \in \{0, 1\}$ governed by the Markov chain appearing in (2.2), i.e., a 2×2 transition matrix potentially different each quarter. The second type of Markov switching-regime model considered differs from the first with regard to the Markov chain dynamics. Indeed, it is assumed that

$$(4.2)\quad p(i_t = i \mid i_{t-1} = i, z_t) = \frac{\exp(z_t' \gamma_i)}{1 + \exp(z_t' \gamma_i)}$$

with i = 0, 1, which is a special case of (2.17) assuming $\gamma_i(\cdot)$ independent of s_t. Of course, it is the choice of the elements entering the vector z_t that is important. Here, we consider two specifications. First, it will be assumed that $z_t = (1, z_{1t})$, where $z_{1t} = (1, \Delta \log GNP^{NSA}_{t-1} - \Delta \log GNP^{SA}_{t-1})$, yielding :

$$(4.3)\ p(i_t = i|i_{t-1} = i, z_t) = \frac{\exp(\gamma_{i0} + \gamma_{i1}(\Delta\log GNP^{NSA}_{t-1} - \Delta\log GNP^{SA}_{t-1}))}{1 + \exp(\gamma_{i0} + \gamma_{i1}(\Delta\log GNP^{NSA}_{t-1} - \Delta\log GNP^{SA}_{t-1}))}.$$

Consequently, when $\gamma_{i1} \neq 0$ for either i = 0 or i = 1, then the *seasonal component* of GNP growth affects the transition probabilities for the seasonally adjusted growth rate in GNP. The hypotheses of interest then become $\gamma_{i1} = 0$ for i = 0 and/or i = 1. Again, when $\gamma_{i1} = 0$ for both i= 0 and i = 1, we recover Hamilton's original model, except that the transition probabilities have been reparameterized to $p = \exp(\gamma_{10}) / (1 + \exp(\gamma_{10}))$ and $q = \exp(\gamma_{00}) / (1 + \exp(\gamma_{00}))$. The second specification we considered involves $z_t = (1, z_{2t})$ where $z_{2t} = (1, (1 - L^2) \log GNP^{NSA}_{t-1} - (1 - L) \log GNP^{SA}_{t-1})$. There is only one difference between this specification and the previous one. As there appears to be fairly strong evidence of a bi-annual unit root in GNP^{NSA}_t, as noted in footnote 14, a removal of such a root is considered. While we did not explicitly discuss ergodicity conditions for a process driven by a Markov chain, as specified in (4.2), it is clear that a nonstationary z_t process would not be acceptable. Hence, by applying (1 + L) and (1 - L) or $(1 - L^2)$ to the log GNP^{NSA}_t, we ensure stationary behavior of the transformed series in the event a bi-annual root is present. Of course, as will be discussed later, it should be noted that the seasonal behavior of (1 - L) log GNP^{NSA}_{t-1} and (1 - L)(1 + L) log GNP^{NSA}_{t-1} is very different.

The left panel of Table 4.2 summarizes the results of what might be called the two most extreme model specifications, one representing the aperiodic model of Hamilton and the other being a periodic Markov chain model with each p and q varying with the season, i.e., involving eight instead of two transition probabilities. Hamilton's original model involved nine parameters, namely four polynomial lag parameters ϕ_1 through ϕ_4, two mean shift parameters α_0 and α_1 defined in (4.1), the innovation variance σ and, finally, the Markov process parameters p and q. The latter two are replaced by p_1 through p_4 and q_1 through q_4 in the periodic model specification. The LM test for periodicity based on the parameter estimates of Hamilton's model as well as the LR test for periodicity are also reported in Table 4.2.

The results in the left panel of Table 4.2 may perhaps be labeled as a mixture of pleasant and unpleasant surprises (we did not report standard errors in the table for the periodic model, an issue that will be dealt with later). First, it should be noted that the parameter estimates other than those pertaining to the Markov chain are statistically, insignificantly different. This means, among other things, that the two states may still be labeled as "recession" and "expansion" states. The surprise is, of course, the multiple boundary parameter estimates. The LM test for periodicity strongly rejects the aperiodic specification. The LR statistic is borderline. If we rely on the standard errors of the aperiodic model, there is also a strong indication that the model is periodic in the sense that the latter parameter estimates are not within two standard errors of the Hamilton model estimates of p and q. We cannot formulate a formal Wald test in this case.[19] In the next section, we shall return to the boundary issue (and, in fact, the right panel of Table 4.2). First, however, we turn to equation (4.3).

Let us now turn to the second model specification appearing in (4.3) with estimates reported in Table 4.3. The nonseasonal model, i.e., with $\gamma_{00} = 2.243$ and $\gamma_{10} = 1.124$ corresponds to a reparameterization of Hamilton's model as $0.904 = \exp(2.243) / (1 + \exp(2.243))$ and $0.755 = \exp(1.124) / (1 + \exp(1.124))$. The same estimates are obtained for the other parameters. We consider first the logistic seasonal model involving z_{1t} and notice that $\hat{\gamma}_{01} = 0.470$ is significant at 3 %, while $\hat{\gamma}_{11} = -0.209$ is not significantly different from zero. The opposite sign of the two

19 What is probably most surprising is the pattern of boundary estimates. In all but the first quarter, we tend to switch to recessions. In all quarters except the first and fourth, we tend to leave expansions. That many boundry estimates is clearly an exageration. Yet, if we let the data speak, only GNP of course, this exageration emerges as a result. It is clear that less of a case of boundary solutions would occur if we were to rely on more series besides GNP, at least that can be implicitly derived from the fact that the NBER chronology for the comparable period of the post-WWII era yields only two boundary parameter estimates, both for recessions (q_1 and q_2). See Ghysels (1990b, Table 5.1).

Table 4.2
Classical ML and Bayes/Gibbs Estimates of Aperiodic and Periodic Markov Switching Regime Models

	Classical ML Estimates of Aperiodic and Unrestricted Periodic Models			Mean and Standard Error of Bayesian Posterior					
	Hamilton Model (Aperiodic)		Unrestricted Periodic Model	Aperiodic Model		Periodic Models			
				1-p ~ Beta (5, 45) 1-q ~ Beta (13,37)		1-p ~ Beta (1, 9) 1-q ~ Beta (6, 18)		1-p ~ Beta (5, 45) 1-q ~ Beta (10, 30)	
	Estimates	Standard Error	Estimates	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
α_1	1.164	0.074	1.178	0.925	0.167	0.968	0.150	0.957	0.152
α_0	-0.359	0.265	-0.192	-0.331	0.244	-0.264	0.201	-0.275	0.199
ϕ_1	0.013	0.120	0.077	0.214	0.082	0.214	0.081	0.203	0.079
ϕ_2	-0.058	0.138	-0.025	0.167	0.077	0.175	0.077	0.167	0.074
ϕ_3	-0.247	0.107	-0.249	-0.102	0.075	-0.097	0.083	-0.101	0.078
ϕ_4	-0.213	0.111	-0.191	-0.121	0.075	-0.121	0.076	-0.123	0.074
σ	0.769	0.103	0.786	0.880	0.063	0.860	0.060	0.871	0.062
p	0.904	0.038	-	0.906	0.036	-	-	-	-
1	-	-	0.684	-	-	0.802	0.122	0.876	0.044
2	-	-	0.999	-	-	0.923	0.075	0.909	0.039
3	-	-	0.999	-	-	0.903	0.084	0.907	0.037
4	-	-	0.999	-	-	0.943	0.055	0.915	0.036
q	0.755	0.097	-	0.722	0.057	-	-	-	-
1	-	-	0.568	-	-	0.711	0.093	0.726	0.070
2	-	-	0.999	-	-	0.754	0.088	0.743	0.065
3	-	-	0.999	-	-	0.768	0.079	0.757	0.067
4	-	-	0.369	-	-	0.732	0.084	0.740	0.069
Log likelihood	-60.88		-55.79						

Notes : There are no standard errors in the unrestricted periodic model. Apart from the Beta distributions, specified in the left panel of the table, all other priors involved in the Bayes/Gibbs estimation setup were taken from McCulloch and Tsay (see Appendix A.2).

Table 4.3
ML Estimation of Logistic Nonseasonal and Seasonal Markov Switching Regime Models

	Nonseasonal Model		Unrestricted Seasonal Model with $z_t = (1, z_{1t})$		Restricted Seasonal Model with $z_t = (1, z_{1t})$		Unrestricted Seasonal Model with $z_t = (1, z_{2t})$		Restricted Seasonal Model with $z_t = (1, z_{2t})$	
	Estimates	Standard Error	Estimates	Standard Error	Estimates	Standard Error	Estimates	Standard Error	Estimates	Standard Error
α_1	1.164	0.074	1.153	0.060	1.152	0.061	1.164	0.080	1.152	0.075
α_0	-0.359	0.265	-0.304	0.135	-0.306	0.141	-0.349	0.211	-0.448	0.261
ϕ_1	0.0135	0.123	-0.039	0.097	-0.039	0.096	0.023	0.128	0.021	0.127
ϕ_2	-0.058	0.138	-0.042	0.099	-0.059	0.097	-0.059	0.127	-0.081	0.143
ϕ_3	-0.247	0.107	-0.313	0.090	-0.303	0.091	-0.231	0.111	-0.248	0.118
ϕ_4	-0.213	0.111	-0.234	0.093	-0.236	0.093	-0.205	0.116	-0.214	0.112
σ	0.769	0.103	0.781	0.085	0.783	0.087	0.601	0.109	0.583	0.109
γ_{10}	2.243**	0.435	3.140**	0.771	3.146**	0.779	1.746**	0.592	2.335**	0.417
γ_{11}	-	-	0.470**	0.222	0.473**	0.225	0.144	0.142	-	-
γ_{00}	1.124**	0.521	1.505**	0.534	1.372**	0.451	0.975*	0.569	0.889	0.577
γ_{01}	-	-	-0.209	0.306	-	-	0.236*	0.143	0.217	0.138
Log likelihood	-60.88		-59.76		-60.00		-58.49		-58.96	
LR test (periodic vs aperiodic)			2.240 0.336		1.760 0.185		4.780* 0.090		3.848** 0.050	

Seasonal means z_{1t} : Q1 = −0.071; Q2 = 0.051; Q3 = 0.004; Q4 = 0.058

Seasonal means z_{2t} : Q1 = 0.013; Q2 = 0.042; Q3 = 0.078; Q4 = 0.082

Notes : * Significant at 10 %; ** significant at 5 % (applicable to Markov chain parameters only).

The model that is estimated appears in (4.3), where $z_{1t} = \Delta \log GNP^{NSA}_{t-1} - \Delta \log GNP^{SA}_{t-1}$ and $z_{2t} = (1 - L^2) \log GNP^{NSA}_{t-1} - \Delta \log GNP^{SA}_{t-1}$, the latter removing the bi-annual unit root in GNP^{NSA}_t.

coefficients is interesting, although one is not significant. Indeed, as the seasonal increases, then the probability of staying in an expansion *increases* while the probability of staying in a recession *decreases*. A priori this result is perfectly plausible, though it should again be stressed that γ_{01} is not significant. The fact that the latter is not significant is not surprising, as there are relatively few observations during recessions, an issue to which we shall return. On the down side, it should be noted that according to the LR test, even if we reestimate the model, only including γ_{11} as the next set of columns (labelled restricted model with z_{1t}) indicates, does not favor the seasonal model as it only yields a p-value of 0.185.

There are two plausible explanations for the conflicting inference drawn from the Wald statistics and the LR tests. As the logistic seasonal model is nonlinear, it is known that the sampling distribution of Wald tests is not as well approximated by the asymptotic distribution as is the case with the LR test [see, e.g., Gallant (1987) or Gallant and White (1988) for further discussion]. This would imply that we reject the seasonal model. Another possibility is that z_{1t} is in fact nonstationary, involving a bi-annual unit root, as previously discussed empirical evidence suggests. In that case, neither the Wald nor the LR tests probably have the correct size asymptotically. To control such a possibility, we consider the specification involving z_{2t}. The drawback of this specification is that one breaks the usual seasonal pattern. Indeed, as reported at the bottom of Table 4.3, we notice that the seasonal means now appear to be low in the first two quarters and high in the second half of the year.[20] With z_{2t} entering the logistic function, the LR test tends to support the seasonal logistic model, while the individual Wald tests do not. The latter, however, is not surprising perhaps as the seasonal pattern of z_{1t}, which was straightforward to interpret, has been reshuffled via the (1 + L) operator for the removal of the bi-annual root. In fact, we find more evidence that γ_{01} is significant, though this disappears with restricted models.

4.3 Bayes estimation of periodic switching-regime models

We now return to the periodic specification reported in Table 4.2. The unconstrained periodic switching-regime model is overparameterized for the amount of

---

20 We report seasonal means both for z_{1t} and z_{2t}, yet it should be noted that when there is a seasonal unit root, the OLS estimates of the seasonal means of z_{1t} are inconsistent and their distribution actually diverges asymptotically (see, for instance, Lee and Siklos (1992) for further discussion).

data available. In a sample of forty years, not enough switches are observed to estimate all eight switching probabilities. Larger data sets would, of course, be more desirable, yet pre-WWII data are not so coherently available. Moreover, the "Great Depression" does not appear to fit easily in the two-state framework of expansion and recession. One possible way to smooth the transition probabilities and hence obtain interior solutions is to adopt a Bayesian estimation strategy. This approach is pursued in this section. It will allow us to tie down the transition probabilities through a common prior, as formulated in (3.16) and (3.17). Indeed, with a Beta prior, one can strike a balance between the desire to smooth the transition probabilities and nevertheless allow for periodic heterogeneity. Preventing transition probabilities from taking extreme values, via smoothing, is a priori reasonable on the basis that irrespective of the season, the economy may always change regime, and hence no season has "absorbing state" features where the economy would be trapped in either one of the regimes. Of course, the choice of the prior is always critical in Bayesian analysis. As smoothing of $p(\mathfrak{s})$ and $q(\mathfrak{s})$ across $\mathfrak{s}$ is critical it will be important to control for the dispersion of the Beta prior. If the Beta prior is very tight, then with 44 years of sample data spread over four quarters and two regimes we may not expect much action in the posterior other than the information in the Beta prior. Hence, a very tight prior induces extreme smoothing and yields essentially on aperiodic models in moderate sample sizes. As the prior is more dispersed, we express more a priori uncertainty about the transition probabilities, though they are tied down by the same prior each quarter. With more uncertainty, heterogeneity in transition probabilities will come more into play. Of course, the extreme of loosening the prior is a complete flat prior, in which case the sample likelihood function determines the shape of the posterior, which, as we know from the previous section, yields corner solutions. Hence, the tightness of the prior will play a crucial role. We considered two prior specifications for $1 - p(\mathfrak{s})$, $\mathfrak{s} = 1, \dots, 4$: namely, we set $a_{11} = 5$ and $a_{12} = 45$ which implies a prior mean for $p(\mathfrak{s}) = 0.89$ for all seasons with the 1st percentile at about 0.80 and the 99th percentile at about 0.97, and $a_{11} = 1$ and $a_{12} = 9$ yielding the same prior mean for $p(\mathfrak{s})$ but the 25th percentile is now at about 0.85, while the 1st percentile is at about 0.60. Consequently, we accept a wider range of possible outcomes each season. This range, however, is not unreasonable, if we consider the parameter estimates appearing in the last column of Table 4.2. Similarly, for $1 - q(\mathfrak{s})$, we set $a_{01} = 10$ and $a_{02} = 30$ yielding a prior mean of 0.67 for $q(\mathfrak{s})$ with a 1st percentile at about 0.60 and $a_{01} = 6$ and $a_{02} = 18$ yielding the same prior mean yet with a larger dispersion. Again, the range is not unreasonable, particularly in the context of the MLE results in Table 4.2.[21]

Table 4.2 summarizes the results from the Gibbs sampler estimation procedure described in section 3.3. A first observation to make is that all parameters other than Markov chain parameters have posterior distributions very much like the ones McCulloch and Tsay (1992) found with their nonperiodic setup which for convenience sake are also reported in the first column of the right panel of Table 4.2.[22] Before discussing the most important aspect, namely the Markov chain parameters, let us briefly point out some of the similarities and differences betwen the MLE results and the means of the Gibbs sampler posteriors. While most of the parameters appear in the same range, it is worth noting that the AR polynomial has a slightly different shape when compared to the MLE estimates obtained by Hamilton.[23] With the loose prior, the estimates of p vary from about 0.80 to about 0.95, i.e., a 15 % range, throughout the year, with the smallest occurring during the winter and the largest during the fall. The fact that the smallest posterior mean occurs during the first quarter should be no surprise, given the pattern of MLE estimates. The standard error of the posterior distribution is also the largest for the first quarter, resulting from the smoothing effect of the prior and the tendency of the likelihood to settle on a much smaller value, namely 0.684. Again, with the relatively loose prior, the means of the posteriors for the recession transition probabilities $q(\cdot)$ range from about 0.71 to about 0.77, or 6 % range. Here, of course, the range is much smaller and closer to the posterior mean of the aperiodic model reported in Table 4.3. The last two columns of Table 4.2 report results with a tight prior. Clearly, the posterior distributions of the transition probabilities cluster more around those of the aperiodic model as would be expected. The moderate size of the sample implies that the information in the likelihood function cannot overturn the strong common prior for p and q.

---

21 We set the priors for the other parameters of the model exactly as in McCulloch and Tsay (1992). Hence, as in the previous section, we followed a similar strategy, namely, we made our setup similar to that of existing estimates of a nonperiodic version of the Markov chain model. Like McCulloch and Tsay, we took a slightly longer sample, though this is inconsequential for all practical purposes. Namely, the data set covers 47:2–91:1, again seasonally adjusted series.

22 This, of course, is very similar to the MLE results reported in Table 4.2 comparing Hamilton's original estimates with those obtained from the periodic model.

23 Indeed, the third- and fourth-order lag coefficients, which as Hamilton noted were relatively large (despite X-11 seasonal adjustment), appear much smaller, though still considerable. Conversely, the first-order lag coefficient is larger now than with MLE estimation and appears "significant" according to the posterior standard errors.

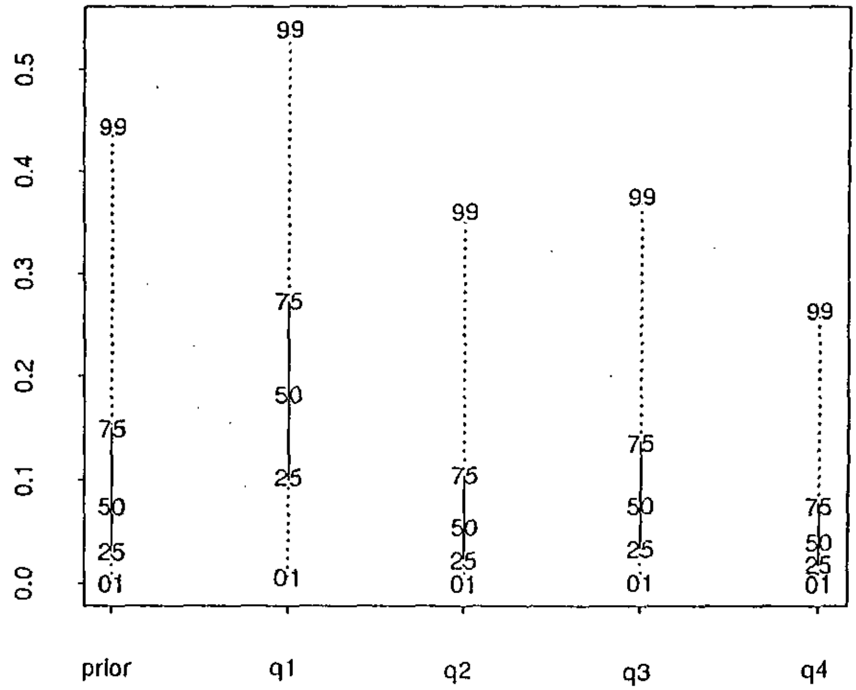

Figure 4.1 : Expansion Switching Probabilities

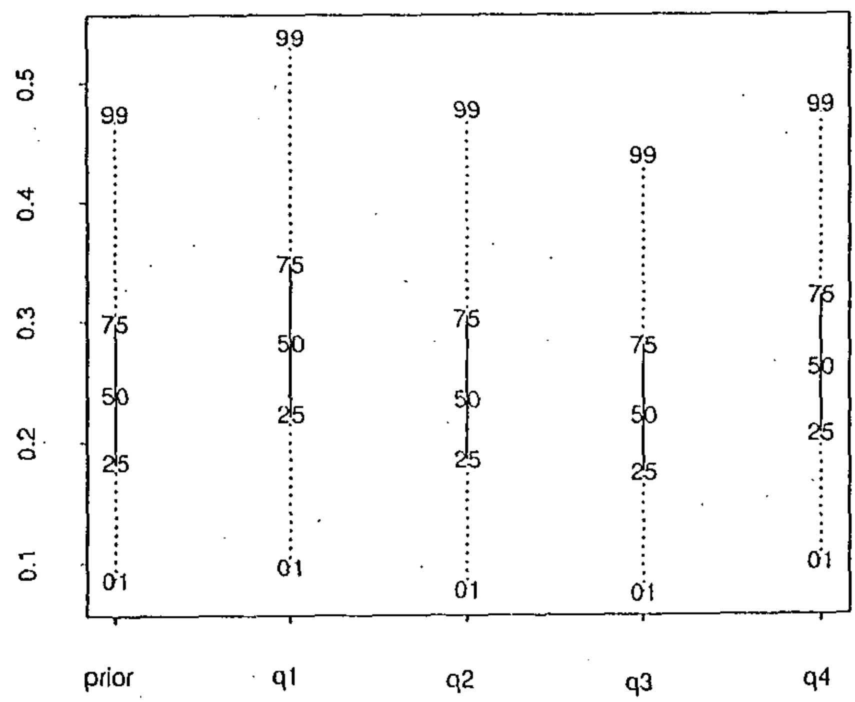

Figure 4.2 : Recession Switching Probabilities

To visualize the contrast between the prior and posterior distributions, let us report our results in two figures. Each figure has five distributions plotted. First, we plot the prior, then we plot the four posterior distributions, one for each quarter. Let us first turn to Figure 4.1. It displays the loose prior for 1 - p(s) and the four posteriors for 1 - p(s), s = 1, ... , 4. First, the fourth quarter very much has a concentrated posterior, relative to the prior, with a higher posterior mean as well. Hence, expansions seem to go through the fall without much risk of staling. We also note that the posterior for the first quarter contains far more evidence now that the probability of staying in an expansion is much lower. The spring and summer show less risk of switching again, yet not as strong as the fourth quarter. As we noted before, the difference between the posterior means p(1) and p(4) is about 15 %, which can be considered as quite significant not only statistically but also economically. Note also that this pattern corresponds to the logistic seasonal model estimates where we found that during "high" seasons, like the fourth quarter, the probability of leaving an expansion decreased. Let us now focus our attention on the recession switching probabilities, appearing in Figure 4.2. Here, the results are far more unsettled. The prior and posterior look very much the same, i.e., the sample information has little impact. There are simply not enough recessions to provide enough sample information, similar to the inprecise estimates of γ_{01} in the logistic seasonal models. There is some movement away from the prior when it is loosely specified, in particular, regarding that the summer has a higher probability of staying in a recession, as the posterior shrinks, relative to the prior. However, the evidence is admittedly not so strong.

5. SEASONAL ADJUSTMENT, FORECASTING AND REGIME-SWITCHING MODELS

One is constantly plagued in econometrics by the perenial questions regarding the purpose, the desirability and the scope of seasonal adjustment. In this section, we will briefly touch on yet another issue regarding seasonal adjustment emerging from the setup of section 2. Then we will move on to practical issues like forecasting turning points of the business cycle with empirical models presented in section 4.

Usually, seasonal characterizations of data are confined to linear structures. The models in section 2 feature nonlinear seasonality, an area relatively unexplored so far. What consequence does this have for the usual approach to seasonal adjustment? In section 4.1, it was noted that some linear features of GNP suggested that periodicity was not fully removed by seasonal adjustment filters. The best way to clarify this is to compute the spectrum of the estimated $\{i_t\}$ process using (2.11). More specifically, we

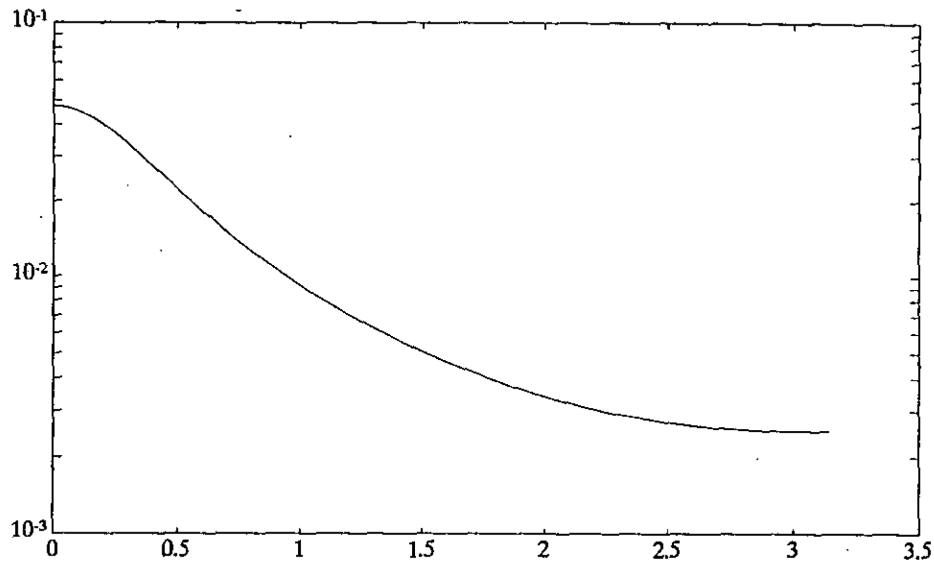

Figure 5.1a : Log Spectrum of Estimated Switching-Regime Process i_t Obtained from the Bayes/Gibbs Estimation

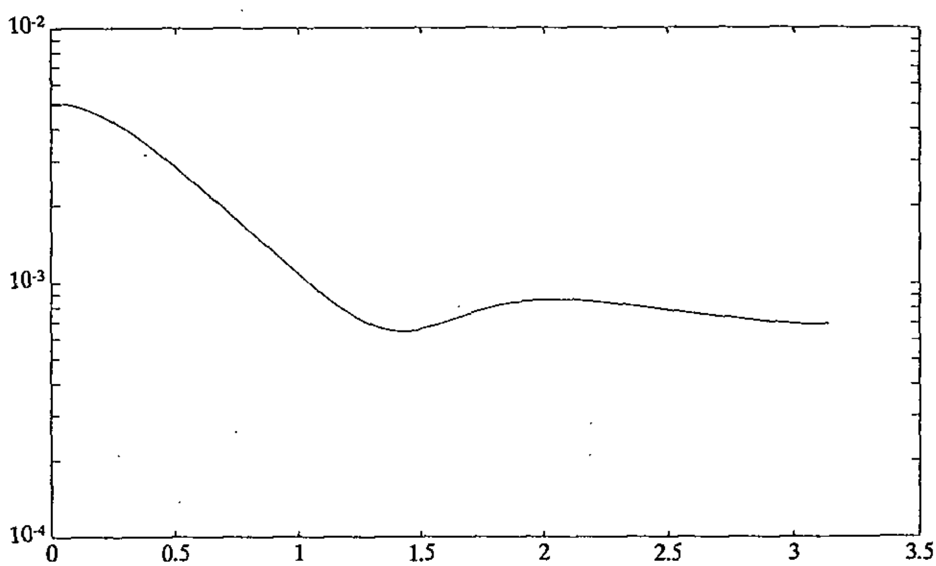

Figure 5.1b : Log Spectrum of Estimated Switching-Regime Process i_t Obtained from the ML Estimation

rely on $\hat{p}(s)$ and $\hat{q}(s)$ in Table 4.2 to characterize the multivariate process in (2.8) and then via Tiao-Grupe's formula obtain its nonperiodic representation. Figure 5.1a displays the spectrum computed from the means of the posterior densities for the switching probabilities, while Figure 5.2b displays those of the MLE estimates (involving the boundary estimates). Clearly, the former has a smooth spectrum, despite differences in p(1) and p(4) up to 0.15. Any usual approach to adjustment would not recognize this as a process with seasonal features. With more extreme values, obtained from MLE, the spectrum displayed in Figure 5.1b shows a dip at the quarterly seasonal frequency. Dips in the spectrum at seasonal frequencies have typically been associated with "overadjustment".

The fact that seasonal adjustment filters do leave traces of periodicity is a fundamental issue, yet a more important question and practically more relevant one regards the forecasting performance of the different models. We consider two types of forecast comparisons, one involving out-of-sample predictions and the other an in-sample comparison of the estimated probabilities that GNP is in a low-growth state.

Table 5.1 reports an out-of-sample performance of four models for seasonally adjusted GNP growth. The four models are : (1) a linear AR(4) model, (2) Hamilton's model, (3) the seasonal logistic model with z_{1t} and (4) the one with z_{2t}. Using the parameter estimates reported in Table 4.3, covering the sample 51:2-84:4, we produced 5 years of one-step-ahead predicitions. For the switching-regime models, we relied on $P[i_t = 0 \mid y_t, ...]$ to calculate the predictions.[24] For the 20 one-step-ahead predicition errors, we calculated mean absolute error and mean squared error measures of accuracy. When we take the linear model as a bench mark, we find that the aperiodic model appears to do worse in predicting GNP growth, while the seasonal models outperform the linear model. Of course, the logistic seasonal models are implicitly multivariate models, since they involve a second series besides GNP growth. Yet, in this particular case, the second series is actually the seasonal component of the series being predicted.

Figure 5.2 reports the inferred probability of the low-growth state over the estimation 53:2 - 84:4 sample obtained from Hamilton's model and the logistic

---

24 This procedure gives a slight advantage to the switching-regime models, in comparison to the linear AR(4) model, by incorporating y_t in the conditioning set.

Table 5.1 : Out-of-Sample Prediction Performance
85:1 - 89:4, Δ Log GNP_t^{SA}
Estimation Sample : 51:2 - 84:4
One-Step-Ahead Prediction Error

	Linear Model	Switching Regime Model	Logistic Seasonal with z_{1t}	Logistic Seasonal with z_{2t}
M.A.E.	0.551	0.626	0.424	0.370
M.S.E.	0.303	0.439	0.255	0.187

seasonal models.[25] The three models are plotted side-by-side. Visual inspection of the plots indicates that the logistic seasonal models appear less erratic, considering, for instance, their behavior during the sixties and mid-seventies. Both are expansion eras where the evidence of seasonal variation was found, which as it appears is taken into account in the logistic seasonal model. Overall, all models track the NBER chronology fairly closely. Besides visual inspection, let us evaluate these models with some formal measures of forecast accuracy. Here, again, we assume that the NBER chronology is the bench mark and compute measures-of-fit for probability forecasts, similar to the analysis of Diebold and Rudebusch (1989). The goodness-of-fit measure is the quadratic probability score (QPS).[26]

From Table 5.2, we learn that with concurrent data (i.e., $k = 0$), the logistic seasonal model with z_{1t} does slightly better than the aperiodic model during NBER recessions, but conversely does worse during NBER expansions. This is not surprising as can be seen in Figure 5.1, since the logistic model has a better coverage of the NBER recessions, particularly during the late fifties, early sixties and early seventies. The logistic seasonal model with z_{2t} does not fare as well, however. It is very interesting to note that all models in Table 5.2 do remarkably worse with $k = 4$ and, in

25 The logistic seasonal models considered are the restricted ones reported in Table 4.3. The NBER chronology used in the quarterly are reported, for instance, in the Appendix of CITIBASE. We did not report the findings for the Bayes/Gibbs model, since they were close to those of the second logistic seasonal model.

26 For exact definitions and further discussion, we refer the reader to Diebold and Rudebusch (1989). It is similar to mean-squared error calculations.

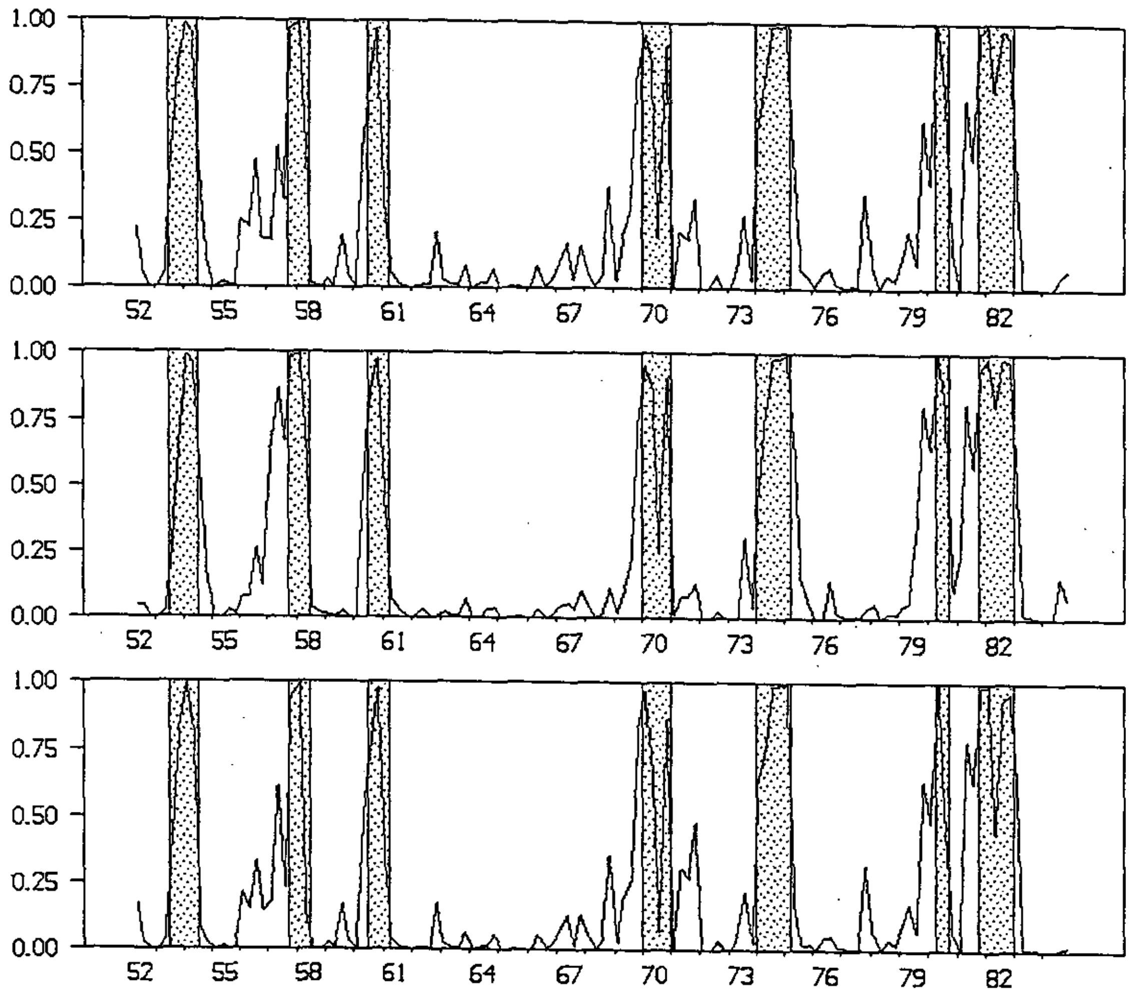

Figure 5.2 : Inferred Probability $P[i_t = 0 \mid y_t, y_{t-1}, \ldots]$
Aperiodic vs Logistic Seasonal Models
(Shaded areas are NBER chronology recessions)
Top : Aperiodic / Middle : Logistic Seasonal with z_{1t} /
Bottom : Logistic Seasonal with z_{2t}

fact, worse than with k = 0. Hence, more data tends to smooth the $P[i_t = 0 \mid y_{t+k} \ldots]$ away from the chronology. How about considering the smoothed probabilities using the entire sample? We notice that all models move closer again to the NBER chronology, yet only one model improves relative to the top panel of Table 5.1, i.e., the concurrent probability assessment. It is the logistic seasonal model with z_{1t}. Its assessment of recessions is improved, while its assessment of expansion eras deteriorates.

Table 5.2
Quadratic Probability Scores for $P[i_t = 0|y_{t+k}, y_{t+k-1}, ...]$ Chronology Against NBER Aperiodic and Logistic Seasonal Models

NBER	Aperiodic	Logistic Seasonal with z_{1t}	Logistic Seasonal with z_{2t}
		k = 0 (Concurrent)	
Dep.	0.158	0.166	0.270
Exp.	0.088	0.118	0.092
		k = 4 (One Year Later)	
Dep.	0.913	0.896	0.994
Exp.	0.426	0.435	0.381
		Smoothed	
Dep.	0.201	0.148	0.255
Exp.	0.166	0.242	0.187

Obviously, this forecasting excercise has the drawback that it relies on the NBER chronology, which is an estimate probably as good, or perhaps worse, than any of the other ones emerging from the models. However, using the chronology as a bench mark, it is clear that using seasonal features, one way or another, appears to improve in some respects the historical record of assessing business cycle phases.[27]

6. CONCLUSIONS

This paper dealt with the possibility of nontrivial interactions between cyclical variation and the repetitive intra-year dynamics of the economy. The presence or absence of such interactions is a fundamental issue. It would be relatively easy to deal with many key issues in macroeconometrics if long spans of uniformly measured time series were available. The parametric structures we presented in this paper lead to straightforward hypotheses one can test regarding periodic features in stochastic regime switching. Like testing for cointegration, long memory, unit roots, mean reversion,

27 The finding that the estimated probablities move away from the NBER chronology may be explained by the fact that the latter relies on the level rather than the growth rate of the series to assess turning points. We made a comparison of the estimated probabilities with the growth-cycle chronology suggested by Moore and Zarnowitz (1986, Table A.8) but found results similar to those reported in Table 5.2.

etc., to name a few key issues, we are hampered by relatively short data sets, like *only* forty or fifty years of data for GNP with not so many regime switches. It is therefore understandable that empirical results are not overwhelmingly conclusive. Yet, the empirical evidence does indicate the presence of interactions between phases of the business cycle and seasonals. For the post-WWII era, the seasonal in GNP growth, with or without taking into account the presence of unit roots at seasonal frequencies, affects regime-switching probabilities in (seasonally adjusted) GNP growth. In terms of forecasting turning points of the economy, there is also evidence of exploiting information contained in the seasonal dynamics. Moreover, a simple assessment of the nonlinear predictability in seasonally adjusted GNP growth also suggests periodicity.

REFERENCES

Albert, J. and S. Chib (1993), "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts", *Journal of Business and Economic Statistics* 11, 1-16.

Bell, W.R. and S. Hillmer (1984), "Issues Involved with the Seasonal Adjustment of Economic Time Series", *Journal of Business and Economic Statistics* 2, 526-534.

Billingsley, P. (1961), "Statistical Inference for Markov Processes" (The University of Chicago Press, Chicago).

Breiman, L. (1969), Probability and Stochastic Processes - with a View Toward Applications (Houghton Mifflin Company, Boston).

Carlin, B.P., A.E. Gelfand and A.F.M. Smith (1992), "Hierarchical Bayes Analysis of Change Point Problems", *Applied Statistics* 41, 389-405.

Carlin, B.P., N. Polson, and D. Stoffer (1992), "A Monte Carlo Approach to Nonnormal and Nonlinear State Space Modelling", *Journal of the American Statistical Association* (forthcoming).

Casella, G. and E.I George (1992), "Explaining the Gibbs Sampler", *The American Statistician* 46, 167-174.

Chib, S. (1991), "Bayes Regression with Autocorrelated Errors : A Gibbs Sampling Approach", *Journal of Econometrics* (forthcoming).

Cox, D.R. and H. Miller (1965), *The Theory of Stochastic Processes* (Chapman and Hall).

Diebold, F.X., J.-H. Lee and G.C. Weinbach (1992), "Regime Switching with Time-Varying Transition Probabilities", mimeo, University of Pennsylvania.

Doob, J.L. (1953), *Stochastic Processes* (John Wiley, New York).

Engel, C. and J.D. Hamilton (1990), "Long Swings in the Dollar : Are They in the Data and Do Markets Know It?", *American Economic Review* 80, 689-713.

Filardo, A.J. (1991), "Business Cycle Phases and Their Transitions", manuscript, Department of Economics, University of Chicago.

Franses, P.H. (1992), "Testing for Seasonality", *Economics Letters* 38, 259-262.

Frisch, R. (1933), "Propagation Problems and Impulse Problems in Dynamic Economics", *Essays in Honor of Gustav Cassel*, George Allen, London.

Gallant, A.R. (1987), "Nonlinear Statistical Models", John Wiley, New York.

Gallant, A.R. and H. White (1987), "A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models" (Basil Blackwell, Oxford).

Garcia, R. and P. Perron (1989), "An Analysis of the Real Interest Rate under Regime Shifts", mimeo, Princeton University.

Gelfand, A.E., S.I. Hills, A. Racine-Poon and A.F.M. Smith (1990), "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling ", *Journal of the American Statistical Association* 85, 972-985.

Gelfand, A.E. and A.F.M. Smith (1990), "Sampling Based Approaches to Calculating Marginal Densities", *Journal of the American Statistical Association* 85, 398-409.

Geman, S. and D. Geman (1984), "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721-741.

Geweke, J. (1991a), "Efficient Simulation from the Multivariate Normal and Student-t Distributrions Subject to Linear Constraints", in Proceedings of the 23rd Symposium on the Interface.

Geweke, J. (1991b), "Bayesian Treatment of the Heteroskedastic Normal and Independent Student-t Linear Models", Working Paper No. 495, Research Department, Federal Reserve Bank of Minneapolis.

Geweke, J. (1992), "Priors for Macroeconomic Time Series and Their Application", Institute for Empirical Macroeconomics Discussion Paper No. 64, Federal Reserve Bank of Minneapolis.

Ghysels, E. (1990a), "On the Economics and Econometrics of Seasonality", in C.A. Sims (ed.) "Advances in Econometrics - Sixth World Congress" (Cambridge University Press) forthcoming.

Ghysels, E. (1990b), "On Seasonal Asymmetries and their Implications for Stochastic and Deterministic Models of Seasonality", mimeo, C.R.D.E., Université de Montréal.

Ghysels, E. (1991a), "On Scoring Asymmetric Periodic Probability Models of Turning Point Forecasts", *Journal of Forecasting* (forthcoming).

Ghysels, E. (1991b), "Are Business Cycle Turning Points Uniformly Distributed Throughout the Year?", Discussion Paper no. 3891, C.R.D.E., Université de Montréal.

Ghysels, E. (1992), "On the Periodic Structure of the Business Cycle", Discussion Paper no. 1028, Cowles Foundation, Yale University.

Ghysels, E. and A. Hall (1992), "Testing Periodicity in some Linear Macroeconomic Models", Discussion paper, C.R.D.E., Université de Montréal.

Ghysels, E., H.S. Lee and P.L. Siklos (1992), "On the (Mis)Specification of Seasonality and Its Consequences : An Empirical Investigation with U.S. Data", *Empirical Economics* (forthcoming).

Ghysels, E., R.E. McCulloch and R.S. Tsay (1992), "Bayesian Inference for a General Class of Periodic Markov Switching Models" (in preparation).

Ghysels, E. and P. Perron (1993), "The Effect of Seasonal Adjustment Filters on Tests for a Unit Root", *Journal of Econometrics* 55, 57-98.

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. (1991a), "A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions", *Journal of Business and Economic Statistics* 9, 27-39.

Hamilton, J.D. (1991b), "Specification Testing in Markov-Switching Time Series Models", mimeo, University of Virginia.

Hamilton, J.D. (1991c), "Estimation, Inference and Forecasting Time Series Subject to Changes in Regime", prepared for C.R. Rao and G.S. Maddala (eds.), *Handbook of Statistics*.

Hamilton, J.D. (1991d), "State-Space Models", in R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics* 4 (forthcoming).

Hansen, B.E. (1991), "The Likelihood Test under Non-Standard Conditions : Testing the Markov Trend Model of GNP".

Hansen, L.P. (1982), "Large Sample Properties of Generalized Methods of Moments Estimators", *Econometrica* 50, 1029-1054.

Hansen, L.P. and T.J. Sargent (1990), "Recursive Linear Models of Dynamic Economies", Manuscript, Hoover Institution, Stanford University.

Hansen, L.P. and T.J. Sargent (1993), "Seasonality and Approximation Errors in Rational Expectations Models", *Journal of Econometrics* 55, 21-56.

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

Hylleberg, S., R.F. Engle, C.W.J. Granger and B.S. Yoo (1990), "Seasonal Integration and Cointegration", *Journal of Econometrics* 44, 215-238.

Karlsen, H.A. (1990), "Doubly Stochastic Vector AR(1) Processes", Chapter 5 in H.A. Karlsen, *A Class of Nonlinear Time Series Models*, Ph.D. Dissertation, Department of Mathematics, University of Bergen.

Lam, P. (1991), "The Hamilton Model with a General Autoregressive Component : Estimation and Comparison with Other Models of Economic Time Series".

Lee, H.S. and P. Siklos (1992), "Seasonality in Time Series : Money-Income Causality in U.S. Data Revisited", Discussion Paper, Wilfrid Laurier University.

McCulloch, R.E. and P.E. Rossi (1992), "Bayesian Inference and Prediction for Mean and Variance Shifts in Autoregressive Time Series", Technical Report, Statistical Research Center, Graduate School of Business, The University of Chicago.

McCulloch, R.E. and R.S. Tsay (1992), "Statistical Inference of Markov Switching Models with Applications to US GNP", Technical Report No. 125, Graduate School of Business, University of Chicago.

Moore, G.M. and V. Zarnowitz (1986), "The Development and Role of the NBER's Business Cycle Chronologies", in R. Gorden (ed.), *The American Business Cycle. Continuity and Change*, Chicago University Press, 735-780.

Nerlove, M., D.M. Grether and J.L. Carvalho (1979), *Analysis of Economic Time Series - a Synthesis*, Academic Press, New York.

Newey, W.K. and K.D. West (1987), "Hypothesis Testing with Efficient Method of Moments Estimation", *International Economic Review* 28, 777-787.

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. and 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.

Priestley, M.B. (1981), Spectral Analysis and Time Series (Academic Press, New York).

Slutsky, C. (1927), "The Summation of Random Causes as the Source of Cyclical Processes", Conjuncture Institute, Moscow.

Tiao, G.C. and M.R. Grupe (1980), "Hidden Periodic Autoregressive-Moving Average Models in Time Series Data", *Biometrika* 67(2), 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.

Zager, S.L. and M.R. Karim (1991), "Generalized Linear Models with Random Effects : A Gibbs Sampling Approach", *Journal of the American Statistical Association* 86, 76-86.

APPENDIX A.1

In this appendix, we discuss the regularity conditions for asymptotic behavior of the vector process y_t appearing in (2.16) under the restriction that the transition matrix $P(s_t, z_t)$ is independent of z_t. We shall introduce a doubly stochastic vector autoregressive representation to conveniently describe the regularity conditions and follow the example of Tjøstheim (1986) and Karlsen (1990).

First, an autoregressive representation for i_t, similar to equation (2.5) yet more general will be the most convenient representation for our purposes. To do this, let us transcribe the definition of i_t using a slightly different notation. Namely, consider the identity matrix of dimension $r^{\ell+1}$ and let the i^{th} column be denoted by e_i for $i = 1, \ldots, r^{\ell+1}$. Then, e_i will represent the state of the world, namely $i_t = i \Leftrightarrow \xi_t = e_i$. Similar to stacking i_t over an entire year as in (2.7a), we can also obtain :

(A.1.1) $\xi_\tau \equiv (\xi'_{\mathcal{S}(\tau-1)+1}, \ldots, \xi'_{\tau\mathcal{S}})'$

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

$$\text{(A.1.2)}\quad \xi_\tau = \begin{bmatrix} I & 0 & & 0 \\ -F_1 & I & & \vdots \\ & & \ddots & \\ & & F_{\mathcal{S}-1} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \ldots & F_{\mathcal{S}} \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & & \vdots \\ 0 & & 0 \end{bmatrix} \xi_{\tau-1} + \chi_\tau$$

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

---

1 The correspondence between F_s and $P(s)$ is relatively simple and can be found, for instance, in Hamilton (1991d, p. 30).

We now define the by now familiar stacked ship-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, namely :

(A.1.3) $\tilde{y}_t \equiv y_t - b_0(i_t, s_t)\, x_{0t}$

(A.1.4) $\tilde{y}_t^{\ell} \equiv [\tilde{y}_t' \dots \tilde{y}_{t-\ell+1}']'$ (A.1.5) $\delta_t^{\ell} \equiv [\delta_t\, O'_{n(\ell-1)x1}]'$

(A.1.6) $\tilde{\underline{y}}_{\tau}^{\ell} \equiv [(\tilde{y}^{\ell}_{\mathscr{S}(\tau-1)+1})', \dots, (\tilde{y}^{\ell}_{\tau\mathscr{S}})']'$ (A.1.7) $\underline{\delta}_{\tau}^{\ell} \equiv [(\delta^{\ell}_{\mathscr{S}(\tau-1)+1})' \dots, (\delta^{\ell}_{\tau\mathscr{S}})']'$

(A.1.8) $B_0^{\ell}(\xi_t, s_t) \equiv [(b_0(i_t, s_t)x_{0t})'\ O'_{n(\ell-1)\times 1}]'$

(A.1.9) $\mathscr{B}_0^{\ell}(\underline{\xi}_t) \equiv [B_0^{\ell}(i_{\mathscr{S}(t-1)+1}, 1)' \dots B_0^{\ell}(i_{\mathscr{S}t}, \mathscr{S})']'$

where the latter two are stacked versions of the intercept process $b_0(i_t, s_t)$. Finally, it is straightforward to also define :

(A.1.10) $y_t^{\ell} \equiv B_0^{\ell}(\xi_t, s_t) + \tilde{y}_t^{\ell}$ (A.1.11) $\underline{y}_{\tau}^{\ell} \equiv \mathscr{B}_0^{\ell}(\underline{\xi}_{\tau}) + \tilde{\underline{y}}_{\tau}^{\ell}$

From (2.16) and the processes defined in (A.1.3) through (A.1.9), we can characterize now a doubly stochastic representation.

(A.1.12) $\tilde{\underline{y}}_{\tau}^{\ell} = \mathscr{B}^{\ell}(\underline{\xi}_{\tau})\, \tilde{\underline{y}}_{\tau-1}^{\ell} + \underline{\delta}_{\tau}^{\ell}$

where
$$\mathscr{B}^{\ell}(\underline{\xi}_{\tau}) \equiv \begin{bmatrix} I_n & 0 & \\ -B^{\ell}(\xi_{\mathscr{S}(\tau-1)+1}, 1) & I_n & \\ 0 & -B^{\ell}(\xi_{\mathscr{S}\tau-1}, \mathscr{S}-1) & I_n \end{bmatrix}^{-1} \begin{bmatrix} 0 & B^{\ell}(\xi_{\mathscr{S}\tau}, \mathscr{S}) \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}$$

with
$$B^{\ell}(\xi_t, s_t) = \begin{bmatrix} b_1(i_t, s_t) \dots \dots b_{\ell}(i_t, s_t) & & \\ I_n & \ddots & \\ 0 & I_n & 0 \end{bmatrix}$$

and $E\, \underline{\delta}_t^{\ell}(\underline{\delta}_t^{\ell})' = I_{\mathscr{S}} \otimes \Lambda^{\ell}$, where $\Lambda^{\ell} = E\, \delta_t^{\ell}(\delta_t^{\ell})'$.

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

The regularity conditions for the existence of a well-defined autocovariance structure for a general periodic Markov switching-regime process can be presented now. Several formulae 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, finally, (3) a theorem then summarizes the main result.

(1) Basic assumptions

Assumption A.1 : The processes $\{\tilde{\underline{y}}_\tau\}$, $\{\underline{\delta}_\tau\}$ and $\{\underline{\xi}_\tau\}$ are defined on a common probability space $(\Omega, \mathcal{A}_{\mathcal{S}}, P_{\mathcal{S}})$.

Assumption A.2 : The process $\{\underline{\xi}_\tau\}$ is a Markov chain which 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 shall denote $r^{\ell+1}$ by K so that the number of states in S equals $\mathcal{S}$K.[2] To proceed with the next assumption, let us define the sigma algebra :

$$\mathcal{A}^{\tau}_{\mathcal{S}y} \equiv \{\underline{y}_u, u \leq \tau\}$$

Assumption A.3 : The matrix functions $B^{\ell}(\cdot)$ and $\mathcal{B}^{\ell}(\cdot)$ appearing in (2.28) 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 $\mathcal{A}^{\tau}_{\mathcal{S}y}$. Likewise, the matrix functions $B_0(\cdot)$ and $\mathcal{B}_0^{\ell}(\cdot)$ are also measurable with regard to same sigma algebra.

---

2 The probability space used in Assumption A.1 is appropriate to deal with stacked skip-sampled vectors where stacking is based on seasons. In particular, $\mathcal{A}_{\mathcal{S}}$ represents a sigma algebra based on sampling events conditional on the seasons they occur in with the associate probability measure $P_{\mathcal{S}}$. The formal discussion presented here includes as special case models which do not involve periodic Markov chains. This is indeed easy to see, simply replace "τ" by "t" in (A.1.5), and replace the probability space $(\Omega, \mathcal{A}_{\mathcal{S}}, P_{\mathcal{S}})$ by $(\Omega, \mathcal{A}, P)$ while the number of states in Assumption A.2 equals $r^{\ell+1}$. See Hansen and Sargent (1990, chap. 10, appendix) for further digressions on the measure theoretic issues involved.

Assumption A.4 : The process $\{\underline{\delta}_\tau\}$ is a martingale difference sequence with regard to $\mathcal{A}^{\tau}_{\mathcal{S}y}$ and $E\,\underline{\delta}_\tau \underline{\delta}_\tau' = I_{\mathcal{S}} \otimes \Lambda^{\ell} < \infty$.

As the estimation will be likelihood-based, we shall in fact assume that $\underline{\delta}_t$ is i.i.d. $N(0, I_{\mathcal{S}} \otimes \Lambda)$ (as we did in equation 2.16). Yet, the analysis in this section may be used to construct a Generalized Methods of Moments estimator [cfr. Hansen (1982)]. Finally, we also assume :

Assumption A.5 : The processes $\{\underline{\delta}_\tau\}$ and $\{\underline{\xi}_\tau\}$ are mutually independent.

*(2) **The covariance structure***

The basic question of interest is under what circumstances are $\{y_t\}$ and its derived processes integrable in quadratic mean, that is to say $\{y_t(\omega)\}$ belongs to the usual Hilbert space $L^2(\Omega, \mathcal{A}, P)$ or $\{\tilde{\underline{y}}_t(\omega)\}$ belongs to $L^2(\Omega, \mathcal{A}_{\mathcal{S}}, P_{\mathcal{S}})$. As all processes have a doubly stochastic representation, we rely on the analysis in Karlsen (1990) to develop necessary conditions for the existence of a well-defined covariance structure.

We shall begin with a discussion of the first moments. Hence, we are interested in the mean of $\{y_t\}$ as it appears in (2.16). This, of course, means we want to analyze the cross-product of the stochastic process $b_0(\cdot)$ and the fixed regressors x_{0t}. From Assumption A.2, we know that the Markov chain process has a unique steady state distribution. For each season $s = 1, \ldots\ldots, \mathcal{S}$, we characterize the steady state distribution as the solution π_s to :

$$\text{(A.1.13)}\quad \pi_s = F_s \pi_s \qquad s = 1, \ldots, \mathcal{S}$$

where the matrices $F_1, \ldots, F_{\mathcal{S}}$ are given in (A.1.2) and several methods can be used to compute π_s [see, e.g., Hamilton (1991d, p. 27)]. Moreover, the steady-state distribution π of the skip-sampled Markov chain ξ_τ can easily be obtained either via computing $\pi = F\pi$ or else :

$$\text{(A.1.14)}\quad \pi = \mathcal{S}^{-1}(\pi_1' \ldots\ldots \pi_{\mathcal{S}}')'$$

Consider now the seasonal sampling of the x_{0t} process, and let its limit be denoted :

$$\text{(A.1.15)} \lim_{T\to\infty} T^{-1} \sum_{\tau=1}^{T} x_{0(\mathscr{S}(\tau-1)+s)} \equiv \bar{x}_0^{s} \qquad s = 1, \dots, \mathscr{S}$$

if x_{0t} is just a constant then of course $\bar{x}_0^{s} \equiv \bar{x}_0 \; \forall s$, yet when x_{0t} includes seasonal dummies then $\bar{x}_0^{s}$ represents a different $n \times 1$ vector each season. Let

$$\text{(A.1.16)} b_0^K(s) \equiv (b_0(1,s)\bar{x}_0^{s}, \dots, b_0(K,s)\bar{x}_0^{s}) \qquad s = 1, \dots, \mathscr{S}$$

be the matrix of all K possible mean shifts each season s, then the mean of y_t conditional on season s is expressed as follows :

$$\text{(A.1.17)} Ey_t \mid s = b_0^K(s)\, \pi_s \qquad s = 1, \dots, \mathscr{S}$$

while the mean unconditional of s is simply $Ey_t \equiv \mathscr{S}^{-1} \sum_{s=1}^{\mathscr{S}} Ey_t \mid s$

Some special cases of (A.1.17) are worth pointing out. For instance, when x_{0t} is just a constant and $b_0^K(s)$ is not a function of s, as for instance is the case in Hamilton (1989) then $Ey_t \mid s$ is simply the cross-product of π_s with b_0^K, i.e., the expected mean shift under steady state distribution π_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. Of course, it was also noted that this seasonal dummy behavior is tightly parameterized as it is entirely determined 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 (A.1.5), as well as the second moments of the $b_0(i_t, s_t)$ process. To streamline its characterization, we rely on the doubly stochastic representation appearing in (A.1.12). We are interested in the following objects :

$$\text{(A.1.18a)} \Gamma(H) = E\, y_\tau^{\ell} (y_{\tau+H}^{\ell})' \qquad H = 0, 1, \dots$$

$$\text{(A.1.18b)} \Gamma_0(H) = E\, (\mathscr{B}_0^{p}(\xi_\tau)\, (\mathscr{B}_0^{\ell}(\xi_{\tau+H}))'$$

(A.1.19a) $\gamma_{s}(h) = E\, y_t^{\ell}(y_{t+h}^{\ell})'$ for $\forall$ t such that $t = \tau(\mathcal{S} - 1) + s$ and h = 0, 1, ...

(A.1.19b) $\gamma_{0s}(h) = E\, B_0^{\ell}(\xi_t, s)\, (B_0^{\ell}(\xi_{t+h}, \tilde{s}))'$ for $\forall$ t such that $t = \tau(\mathcal{S} - 1) + s$,

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

The formula in (A.1.18) represents the covariance structure for the stacked skip-sampled vector process y_τ^{ℓ}. In contrast, the formula in (A.1.19) represents the covariance structure, conditional on a particular season, of the nonstacked y_t^{ℓ} process. Once the formulae in (A.1.18) and (A.1.19) are well defined and characterized, one can again invoke the Tiao-Grupe formula appearing in (2.11), this time applied to the y_t^{ℓ} process, yielding expressions for :

(A.1.20a) $\gamma(h) = E\, y_t^{\ell}(y_{t+h}^{\ell})'$ $\quad\forall$ t and h = 0, 1, ...

(A.1.20b) $\gamma_0(h) = E(B_0^{\ell}(\xi_t, \cdot)\, (B_0^{\ell}(\xi_{t+h}, \cdot)')$ $\forall$ t and h = 0, 1, ...

The existence and characterization of (A.1.18) through (A.1.20) is determined as follows :

Theorem A.1 : Let Assumptions A.1 through A.5 hold. Then stochastic processes $\{y_\tau^{\ell}\}$ and $\{B_0^{\ell}(\xi_\tau)\}$ are covariance-stationary with $\Gamma(0)$ and $\Gamma_0(0)$ finite if :

(a) $\underset{1 \le s \le \mathcal{S}}{\text{Max}}\ (\Pi'_s \otimes I_{(n\ell)^2})\, [\text{diag}(B_0^{\ell}(k, s)\, B_0^{\ell}(k, s)')_{k=1}^{K}]\, (1_K \otimes I_{(n\ell)^2}) < \infty$

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

where $\rho(\cdot)$ is the spectral radius while 1_j and I_j are respectively a $1 \times j$ vector of ones and an identity matrix of dimension j. The matrix $\mathcal{P}$ is defined in Assumption A.2. Moreover,

(A.1.21) $\text{vec}\,\Gamma(H) = (1_{\mathcal{S}K} \otimes I_{(\mathcal{S}n\ell)^2})\, \mathcal{F}_1^{H}(I_{\mathcal{S}K(\mathcal{S}n\ell)^2} - \mathcal{F}_2)^{-1}\, \mathcal{F}_0\, \text{vec}\, I_{\mathcal{S}} \otimes \Lambda^{\ell}$

$$\text{(A.1.22)}\ \text{vec}\ \Gamma_0(H) = \sum_{k=1}^{\mathscr{S}K} \pi_k \sum_{j=1}^{K} (\tilde{p}^H_{kj} - \pi_j)\, B^{\ell}_0(k)\, B^{\ell}_0(k)'$$

where $\tilde{p}^H_{kj}$ is the kj-th element of $\mathscr{P}^H$ and where the matrices $\mathscr{F}_i$, i = 0, 1 and 2 are as follows :

$$\mathscr{F}_0 = \pi \otimes I_{(\mathscr{s}n\ell)^2}$$

$$\mathscr{F}_1 = [\text{diag}(I_{\mathscr{s}n\ell} \otimes \mathscr{B}^{\ell}(k))_{k=1}^{\mathscr{s}K}](\pi \otimes 1'_{\mathscr{s}K})' \otimes I_{(\mathscr{s}n\ell)^2}$$

$$\mathscr{F}_2 = [\text{diag}(\mathscr{B}^{\ell}(k) \otimes \mathscr{B}^{\ell}(k))_{k=1}^{\mathscr{s}K}](\pi \otimes 1'_{\mathscr{s}K})' \otimes I_{(\mathscr{s}n\ell)}\,.$$

Finally,

$$\text{(A.1.23)}\ \Gamma(H) = \begin{bmatrix} \gamma_1(\mathscr{S}) & \gamma_1(\mathscr{S}H - \mathscr{S} + 1) \\ \vdots & \\ \gamma_{\mathscr{S}}(\mathscr{S}H - \mathscr{S} + 1) & \gamma_{\mathscr{S}}(\mathscr{S}H) \end{bmatrix}$$

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

$$\text{(A.1.24)}\ \text{vec}\ \gamma_{\mathscr{s}}(h) = \pi'_{\tilde{\mathscr{s}}} \otimes I_{(n\ell)^2}[\text{diag}(B^{\ell}(k, \tilde{\mathscr{s}}) \otimes B^{\ell}(k, \tilde{\mathscr{s}})')_{k=1}^{K}]\ \text{vec}\ \gamma_{\mathscr{s}}(h-1)$$

$$+\ \text{vec}\ \Lambda^{\ell}\,.$$

Proof : By formulation equation (2.16) as a doubly stochastic vector AR(1) process, we can use Theorem 4.1 of Karlsen (1990) spelling 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. Condition (a) applies to the $\{B^{\ell}_0\}$ process second-moment structure, while condition (b) applies to the de-measured process $\tilde{y}^{\ell}_{\tau}$. The bounded spectral radius condition follows from (4.2) in Theorem 4.1 of Karlsen (1990). Likewise (A.1.21) and (A.1.22) follow from formulae (4.3) and (4.5). Equation (A.1.23), establishing the relationship between $\Gamma(\cdot)$ and $\gamma_{\mathscr{s}}(\cdot)$, follows from from the stacked skip-sampling structure of $\{y^{\ell}_{\tau}\}$. Finally, equation (A.1.24) follows from the doubly stochastic representation of the same process, without stacking, however.

APPENDIX A.2

In this appendix, we briefly summarize the basic features of the Gibbs sampler used in the paper. As noted in the main body of the paper, our approach closely follows that of McCulloch and Tsay (1992).

We shall first discuss the formulation of the prior as well as the conditional posterior distributions which are easy to obtain from traditional Bayesian analysis. Next, we discuss how to compute the joint posterior distributions.

Let us consider again equation (2.16). For the purpose of formulating the prior, let us assume that :

$$\text{(A.2.1)}\quad b_0(i_t, s_t)\, x_{0t} \equiv b_0^c\, x_{0t} + b_0^u(i_t, s_t)\, x_{0t} + b_0^g(i_t, s_t)$$

where b_0^c is common to all states, b_0^u in contrast is state-specific yet unconstrained while b_0^g is also state-specific but subject to inequality constraints like $b_0^g(i, s) < b_0^g(j, \tilde{s})$ $\forall s, \tilde{s}$ and i and j given with $i \neq j$. This yields the following priors on the parameters governing the mean-shift function of the $\{y_t\}$ process :

$$\text{(A.2.2)}\quad b_0^c \sim N(b_m^c, A_c^{-1})$$

$$\text{(A.2.3)}\quad \mathrm{vec}\begin{bmatrix} b_0^u(1,1) \ldots\ldots b_0^u(1, \mathcal{S}) \\ \vdots \qquad\qquad \vdots \\ b_0^u(K,1) \ldots\ldots b_0^u(K, \mathcal{S}) \end{bmatrix} \sim N(b_m^u, A_u^{-1})$$

$$\text{(A.2.4)}\quad \mathrm{vec}\begin{bmatrix} b_0^g(1,1) \ldots\ldots b_0^g(1, \mathcal{S}) \\ \vdots \qquad\qquad \vdots \\ b_0^g(K,1) \ldots\ldots b_0^g(K, \mathcal{S}) \end{bmatrix} \sim N(b_m^g, A_g^{-1})\, I(\eta_g)$$

where b_m^c, b_m^u, b_m^g are hyper-parameter vectors of dimension $n \times 1$ and twice $(K\ \mathcal{S}\ n) \times 1$ respectively describing the prior mean. Likewise A_c, A_u and A_g are hyper-parameter matrices describing the dispersion of the prior. Finally, $I(\eta_g)$ is a $(K\ s\ n) \times 1$ indicator function whose elements equal 1 if the imposed inequality constraints are satisfied and zero otherwise. The prior for the inverse of the innovation variance Λ^{-1} is assumed to be Wishart distributed $W(\Lambda^{-1}, \nu)$, which in the univariate case, i.e., n = 1 implies $\sigma^2 \sim \nu\,\overline{\sigma}/\chi^2_\nu$ where $\overline{\sigma}$ and ν are again hyper-parameters. The polynomial lag $b_j(i_t, s_t)$ for j = 1, ..., ℓ appearing in (2.16) are assumed independent of (i_t, s_t) and governed by the prior distribution :

$$\text{(A.2.5)}\ \mathrm{vec}(b_1, \dots b_\ell) \sim N(b_m, A_\ell^{-1})$$

again controlled by a set of hyper-parameters. Finally, and more importantly, we consider the prior for the parameters governing the periodic Markov chain, as discussed in the main body of the paper. We confine our attention again to a two-state process, i.e., r = 2, then adopting the notation of (2.2), the priors are formulated as follows :

$$\text{(A.2.6)}\ 1 - 1(s) \sim \mathrm{Beta}(a_{01}, a_{02}) \qquad s = 1, .., \mathcal{S}$$

$$\text{(A.2.7)}\ 1 - p(s) \sim \mathrm{Beta}(a_{11}, a_{12}) \qquad s = 1, .., \mathcal{S}$$

From the aforementioned priors, one can compute conditional posterior distributions which will be used for the Gibbs simulations. McCulloch and Tsay (1992) describe in detail how to compute the following conditional posterior distributions : $p(x \mid Y_{\mathcal{S}T}, [\theta \backslash \{x\}])$ where $x = b_0^c$, $\mathrm{vec}(b_1, \dots, b_\ell)$ and Λ while $[\theta \backslash \{x\}]$ represents the parameter vector θ excluding the elements appearing in x. The former two conditional posteriors are normal, the latter is inverted Wishart. Moreover, let $I_t \equiv \{i_{t-j}, s_{t-j}),\ t > j \geq 1\}$ pretending as if states were observable, then using similar arguments, one can also compute the conditional posterior densities $p(x \mid Y_{\mathcal{S}T}, I_{\mathcal{S}T}\ [\theta \backslash \{x\}])$ where $x = b_0^u$ and b_0^g, the former being a normal distribution while the latter is truncated normal because of the inequality constraints appearing in (A.2.4). Finally, it is worth discussing briefly the conditional posterior distributions for $q(s)$ and $p(s)$. They are :

$$p(p(s) \mid I_{\mathcal{S}T}) \sim \mathrm{Beta}(a_{01}, ef_{01}(s), a_{02}, ef_{00}(s)) \qquad s = 1, .., \mathcal{S}$$

$$p(q(s) \mid I_{\mathcal{S}T}) \sim \mathrm{Beta}(a_{11}, ef_{10}(s), a_{12}, ef_{11}(s)) \qquad s = 1, .., \mathcal{S}$$

where $ef_{ij}(s)$ is the number of "jumps" from i to j in season s given the history of states described by $I_{\mathcal{S}T}$. Note that unlike the classical MLE, the conditional posterior mean will not be at the boundary in any given finite sample even with $ef_{ij}(s) = 0$ for any given i, j and s.

Of course, we are not directly interested in the conditional posterior distributions, but instead the analytically intractable joint posterior for the entire set of parameters as well as the history of unobserved states $I_{\mathcal{T}}$. At this stage, we rely on the Gibbs sampling principle as a simulation tool. Since this technique is now widely used and well documented, notably in the survey paper by Casella and George (1992), we refrain from repeating the iterative steps of computing a conditional posterior distribution. The outcome, provided mild regularity conditions are satisfied [see, e.g., Geman and Geman (1984)] is the joint posterior density of interest.