Instituto Complutense de Análisis Económico
UNIVERSIDAD COMPLUTENSE
Campus de Somosaguas
28223 MADRID

Documento de Trabajo

A General Test for Univariate Seasonality

Rafael Flores
Alfonso Novales

No.9503 (Revisado) Octubre 1995

Instituto Complutense de Análisis Económico
UNIVERSIDAD COMPLUTENSE

A GENERAL TEST FOR UNIVARIATE SEASONALITY

Rafael Flores*
Alfonso Novales*
Instituto Complutense de Análisis Económico
Universidad Complutense
Campus de Somosaguas
28223 Madrid

ABSTRACT

We propose a general test for univariate seasonality. Starting from a multivariate model for the seasons, some constraints must hold both, on the covariance matrix of the innovations, as well as among coefficients across equations, for a univariate representation of seasonality to be appropriate. Appied to a set of 23 U.K. macroeconomic variables, our test shows that a multivariate representation of seasonality should be preferred in at least 8 cases. That introduces a serious questioning of standard, univariate filters to estimate the seasonal component in some economic time series, and suggest the possibility of a more complex, but richer way of characterizing relationships among seasonal economic variables.

RESUMEN

Se propone un contraste para la detección de estacionalidad univariante. Para que la estacionalidad presente en una serie temporal pueda ser modelizada en un contexto univariante, el modelo multivariante estocástico de las estaciones deberá incorporar determinadas restricciones, tanto sobre la matriz de varianzas y covarianzas de las innovaciones como entre los coeficientes de las distintas ecuaciones. Cuando el contraste propuesto se aplica a un conjunto de 23 variables de la economía del Reino Unido, en al menos 8 casos se detecta una estructura multivariante de la estacionalidad. Este hecho introduce serias dudas acerca de los filtros univariantes, estándar, utilizados para la estimación del componente estacional en algunas series económicas. Al mismo tiempo esta evidencia empírica sugiere una forma más compleja pero más rica de caracterizar las relaciones entre las variables económicas.

Acknowledgements

The authors have benefited from discussions with participants at the European Econometric Society meeting at Maastricht (july 1994) and numerous seminars and, in particular, from conversations with J.J. Dolado, P.H. franses and S. Hyllberg.

Key words: Seasonality, PAR models. **JEL Classification:** C32, C52

* Departamento de Economía Cuantitativa. Facultad de Ciencias Económicas y Empresariales. Universidad Complutense. MADRID 28223 (SPAIN). Ph:(91)394-2370/2594. Fax:(91)394-2613. E-mail for R. Flores: eccua04@sis.ucm.es. E-mail for A. Novales: eccua09@sis.ucm.es.

1. INTRODUCTION

The seasonal characteristics inherent to frequently observed economic variables are usually represented within a univariate, stochastic framework. Nevertheless, it is far from obvious that a single source of randomness is enough to capture the seasonal characteristics of economic variables. To the contrary, it seems quite natural to consider that each season may experience a random perturbation different from those for the other seasons.

There might be variables for which the seasonal shocks are very similar to each other, a single perturbation being then able to capture most of the statistical features of that variable. But even if the seasonal shocks were reasonably similar, that would not be enough to validate a univariate representation. The behavior of a given season over the years, as well as the way it relates to previous and subsequent seasons, should also be sufficiently similar across them [see for example, Osborn et al. (1988), Birchenhall et al. (1989), Osborn (1991), Franses and Romijn (1993) and Franses and Paap (1994)].

We propose in this paper a general test for univariate seasonality. The starting point is a VAR model for the s seasons of a given variable. If seasonality was of a univariate nature, some constraints on the contemporaneous residual covariance matrix as well as on the VAR coefficients should hold. Our test considers both sets of constraints. First, the orthogonalized residuals of the s-variate model should have the same variance. Secondly, the coefficients in a given season equation should be equal to those in the other equations.

When applied to a set of 23 U.K. macroeconomic variables, our test shows that a multivariate representation of seasonality should be preferred in some cases. That introduces a serious questioning of the generalized use of standard, univariate filters to estimate the seasonal component in economic time series, and suggests the possibility of a more complex, but richer way of characterizing relationships among seasonal economic variables.

2. MULTIVARIATE VS. UNIVARIATE REPRESENTATIONS OF SEASONALITY

2.1 A Multivariate Representation of Seasonality.

Let x_{ti} be an economic variable being observed at seasonal frequencies. We denote by a time index t the year and by a second index $i=1, \ldots, s$ the period of the year (a month or a quarter) to which the observation corresponds. Typically, $s = 4$ or 12. Let $\mathbf{X}_t$ denote the vector of annual observations: $\mathbf{X}_t' = (x_{t1} \ldots x_{ts})$ for year $t=1,2,\ldots,n$, and let its finite order VAR representation[1] be:

$$\mathbf{\Phi(B)}\ \mathbf{X}_t = \boldsymbol{\mu} + \mathbf{a}_t \qquad (1)$$

where $\mathbf{\Phi(B)}$ is an sxs matrix of polynomials in the lag operator B with $\mathbf{\Phi(0)} = \mathbf{I}_s$ and all the roots of its determinant characteristic equation $|\mathbf{\Phi(B)}| = 0$ being on or outside the unit circle; $\mathbf{a}_t$ is an i., i.d., zero mean, s-variate Gaussian random vector: $\mathbf{a}'_t = (a_{t1}\ a_{t2} \ldots a_{ts})$ with contemporaneous covariance matrix $\mathbf{\Sigma}$, representing the cross-effects among seasons in the same year. Differences among the elements in μ can be thought of as being related to the deterministic component of seasonality in $\mathbf{X}_t$, although we do not pursue that line of reasoning here. For simplicity of exposition, we will ignore in what follows this vector of constants, although it will be included when estimating model (1) in section 3.

Having introduced a seasonal variable as the realization of a multivariate vector, it is now of interest to test whether for some variables, the s-variate structure reduces to a univariate one. In section 3 we propose a likelihood ratio test statistic based on the orthogonalized residuals from VAR model (1) fitted to the s series for the seasons.

The Hillmer and Tiao(1979) representation of a VAR(p) model for the entire sample is:

$$\mathbf{D}_\phi\ \mathbf{X} = \mathbf{a} - \mathbf{G}_\phi\ \mathbf{X}^* \qquad (2)$$

where $\mathbf{X}$ denotes the ns-vector of sample observations $\mathbf{X}' = (\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_n')$, with n being the number of years, $\mathbf{a}' = (\mathbf{a}_1', \mathbf{a}_2', \ldots, \mathbf{a}_n')$ is the $ns\times 1$ noise vector with a non scalar variance-covariance matrix $(\mathbf{I}_n \otimes \mathbf{\Sigma})$, $\mathbf{\Sigma}$ being the $s\times s$ intra-year variance-covariance

matrix, and $\mathbf{X}^*$ is the $ps\times 1$ vector of initial conditions. $\mathbf{D}_\phi$ is the $ns\times ns$ matrix of parameters:

$$\mathbf{D}_\phi = \begin{bmatrix} \mathbf{I}_s & \mathbf{0} & \mathbf{0} & .. & \mathbf{0} & \mathbf{0} & .. & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{\phi}_1 & \mathbf{I}_s & \mathbf{0} & .. & \mathbf{0} & \mathbf{0} & .. & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{\phi}_2 & -\boldsymbol{\phi}_1 & \mathbf{I}_s & .. & \mathbf{0} & \mathbf{0} & .. & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ -\boldsymbol{\phi}_p & -\boldsymbol{\phi}_{p-1} & -\boldsymbol{\phi}_{p-2} & .. & \mathbf{I}_s & \mathbf{0} & .. & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\phi}_p & -\boldsymbol{\phi}_{p-1} & .. & -\boldsymbol{\phi}_1 & \mathbf{I}_s & .. & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & .. & -\boldsymbol{\phi}_{p-1} & -\boldsymbol{\phi}_{p-2} & .. & \mathbf{I}_s & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & .. & -\boldsymbol{\phi}_p & -\boldsymbol{\phi}_{p-1} & .. & -\boldsymbol{\phi}_1 & \mathbf{I}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & .. & \mathbf{0} & -\boldsymbol{\phi}_p & .. & -\boldsymbol{\phi}_2 & -\boldsymbol{\phi}_1 & \mathbf{I}_s \end{bmatrix} \qquad (3)$$

where $\mathbf{I}_s$ is the $s\times s$ identity matrix and $\boldsymbol{\phi}_i$ $(i=1, \ldots, p)$ is the $s\times s$ matrix of parameters associated with lag i in VAR(p) model (1).

In (2), $\mathbf{G}_\phi$ is the $ns\times ps$ matrix of parameters:

$$\mathbf{G}_\phi = \begin{bmatrix} -\boldsymbol{\phi}_p & -\boldsymbol{\phi}_{p-1} & .. & .. & -\boldsymbol{\phi}_2 & -\boldsymbol{\phi}_1 \\ \mathbf{0} & -\boldsymbol{\phi}_p & .. & .. & -\boldsymbol{\phi}_3 & -\boldsymbol{\phi}_2 \\ .. & .. & .. & .. & .. & .. \\ \mathbf{0} & \mathbf{0} & .. & .. & -\boldsymbol{\phi}_p & -\boldsymbol{\phi}_{p-1} \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & -\boldsymbol{\phi}_p \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & \mathbf{0} \\ .. & .. & .. & .. & .. & .. \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad (4)$$

Let $\mathbf{T}$ be a $s\times s$ lower triangular matrix, constrained to have ones in the main diagonal, such that $\mathbf{T\Sigma T'}=\mathbf{\Lambda}$ is a diagonal matrix. Premultiplying (2) times $(\mathbf{I}_n\otimes\mathbf{T})$ we get:

$$(\mathbf{I}_n\otimes\mathbf{T})\ \mathbf{D}_\phi\ \mathbf{X} = (\mathbf{I}_n\otimes\mathbf{T})\ \mathbf{a} - (\mathbf{I}_n\otimes\mathbf{T})\ \mathbf{G}_\phi\ \mathbf{X}^* \qquad (5)$$

or:

$$\mathbf{D}_\phi^u \mathbf{X} = \epsilon - \mathbf{G}_\phi^u \mathbf{X}^* \tag{6}$$

where:

$$\begin{aligned} \mathbf{D}_\phi^u &= (\mathbf{I}_n \otimes \mathbf{T})\, \mathbf{D}_\phi \\ \mathbf{G}_\phi^u &= (\mathbf{I}_n \otimes \mathbf{T})\, \mathbf{G}_\phi \\ \epsilon &= (\mathbf{I}_n \otimes \mathbf{T})\, \mathbf{a} \end{aligned} \tag{7}$$

with: $Var(\epsilon) = (\mathbf{I}_n \otimes \Lambda)$. We will refer to (6) as the orthogonalized VAR.

So long as Σ is positive definite, matrix **T** can be obtained as the factor in the Choleski decomposition [see Theil (1983), p.34, for instance]. For that, we project the residual for a given season from a previously estimated VAR on the residuals corresponding to the previous quarters of that same year. Ordering the obtained coefficient estimates in each of these projections as one of the rows below the main diagonal produces a matrix **T** with the above mentioned properties. Each element in ϵ is the residual in these auxiliary regressions. Residuals corresponding to the first season remain unchanged relative to vector **a**. Proceeding this way, we are interpreting the covariances in Σ as measuring the influence of previous seasons on the current one.

Expression (6) reduces to the Hillmer-Tiao representation for a univariate AR(r) process, $r \leq sp$, when: (a) the elements in any main lower diagonal of $\mathbf{D}_\phi^u$ are all equal to each other and (b) Λ is a scalar matrix. Thus, lack of heteroskedasticity is a necessary but not sufficient condition for seasonality to have a univariate representation. If conditions a) and b) both hold, then a standard, univariate modelling approach, may be appropriate.

For instance, in the case of a VAR(2) with s=4 and matrices ϕ_1, ϕ_2, then $\mathbf{D}_\phi^u$ takes the general form:

$$\mathbf{D}_\phi^u = \begin{bmatrix} \mathbf{T} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}\phi_1 & \mathbf{T} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}\phi_2 & -\mathbf{T}\phi_1 & \mathbf{T} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{T}\phi_2 & -\mathbf{T}\phi_1 & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{T} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{T}\phi_1 & \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{T}\phi_2 & -\mathbf{T}\phi_1 & \mathbf{T} \end{bmatrix} \tag{8}$$

with **T** and ϕ_k, k=1,2, being 4x4 matrices of parameters.

If the data generating process was indeed univariate, the matrices in $\mathbf{D}_\phi^u$ would have a structure:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \pi_{14,1} & 1 & 0 & 0 \\ \pi_{13,1} & \pi_{14,1} & 1 & 0 \\ \pi_{12,1} & \pi_{13,1} & \pi_{14,1} & 1 \end{bmatrix} \tag{9}$$

$$\mathbf{T}\phi_1 = \begin{bmatrix} \pi_{11,1} & \pi_{12,1} & \pi_{13,1} & \pi_{14,1} \\ \pi_{14,2} & \pi_{11,1} & \pi_{12,1} & \pi_{13,1} \\ \pi_{13,2} & \pi_{14,2} & \pi_{11,1} & \pi_{12,1} \\ \pi_{12,2} & \pi_{13,2} & \pi_{14,2} & \pi_{11,1} \end{bmatrix} \qquad \mathbf{T}\phi_2 = \begin{bmatrix} \pi_{11,2} & \pi_{12,2} & \pi_{13,2} & \pi_{14,2} \\ 0 & \pi_{11,2} & \pi_{12,2} & \pi_{13,2} \\ 0 & 0 & \pi_{11,2} & \pi_{12,2} \\ 0 & 0 & 0 & \pi_{11,2} \end{bmatrix} \tag{10}$$

where $\pi_{ij,k}$ represents the element (i,j) of matrix $\mathbf{T}\phi_k$. Excluding the vector of constants, there would be just 8 different parameters, out of the initial 32 coefficients in (1).

The associated univariate autoregressive process would be at most an AR(8)[2]:

$$(1+\pi_{14,1}B+\pi_{13,1}B^2+\pi_{12,1}B^3-\pi_{11,1}B^4-\pi_{14,2}B^5-\pi_{13,2}B^6-\pi_{12,2}B^7-\pi_{11,2}B^8)\, x_j = \epsilon_j \tag{11}$$

j = 1,2,...,ns, where we now use subindex j to denote each one of the observations in the ns-vector **X**.

For the VAR(1) with s=4 case, the matrices would be:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \pi_{14,1} & 1 & 0 & 0 \\ \pi_{13,1} & \pi_{14,1} & 1 & 0 \\ \pi_{12,1} & \pi_{13,1} & \pi_{14,1} & 1 \end{bmatrix} \qquad \mathbf{T}\phi_1 = \begin{bmatrix} \pi_{11,1} & \pi_{12,1} & \pi_{13,1} & \pi_{14,1} \\ 0 & \pi_{11,1} & \pi_{12,1} & \pi_{13,1} \\ 0 & 0 & \pi_{11,1} & \pi_{12,1} \\ 0 & 0 & 0 & \pi_{11,1} \end{bmatrix} \tag{12}$$

with just 4 different coefficients in (1) out of the starting 16, and the equivalent univariate AR(4)[3] process wuld be:

$$(1+\pi_{14,1}B+\pi_{13,1}B^2+\pi_{12,1}B^3-\pi_{11,1}B^4)\, x_j = \epsilon_j, \qquad j = 1,2,...,ns \tag{13}$$

2.2 Univariate Representations of Seasonality.

For comparison with previous research on these issues, we use the same data set as in Osborn (1990), which has recently been considered by Franses and Romijn (1993). It includes quarterly variables, in logs, except by the yield on Treasury bills, and the physical increase in stocks. Trade balance is computed as the difference between log exports and log imports[4]. A list of abbreviations for the names of the variables is provided at the end of the paper.

Table 1 presents our estimates of the univariate models specified for the variables in our sample with the differencing order[5] suggested in Osborn (1990). Pure autoregressions seem to be adequate for all variables. They are characterized by a small number of parameters, to a maximum of six, and represent seasonality with a second order multiplicative autoregressive term in most cases. The Lunj-Box statistic suggests possible misspecification, but this is due to the presence of outliers, on which we will comment in section 3.2. Alternative mixed ARMA representations do not greatly reduce the number of parameters with respect to the AR models, being the fit of data in both cases very similar. The reader is refered to Table A2.1 in appendix 2 for more details about the estimated mixed models.

To the effect of comparing univariate with multivariate representations of seasonality, we consider in what follows the AR models estimated under the differencing order in Osborn (1990).

[INTRODUCE TABLE 1 HERE]

2.3 Choosing the Lag Order in the VAR Representation for the Seasons.

Choosing the lag length for the VAR representation for the seasons should be expected to condition the results. Two approaches can be considered in our framework: on the one hand, an estimated univariate model implies a minimum order for a VAR for the seasons to be consistent with it. On the other, we can use standard likelihood ratio test to select the lag length of the VAR.

If the univariate models are taken to be correct, then the corresponding VAR representation for the seasons should be of relatively high order. To accomodate our estimated univariate models, a third or fourth order VAR is needed for 21 out of the 23

variables considered. Only for the financial variables: the Financial Times index of industrial ordinary shares and Treasury Bill interest rates, a VAR(1) would be enough [see Table 2, column 3].

Using the likelihood ratio test for the VAR order, we incorporated the correction in Sims (1980) by substracting from the number of years **n** the number **c** of estimated coefficients in each equation of the unrestricted VAR. The statistic was then computed as the product of **n-c** times the difference between the logarithms of the determinants of the covariance matrices of the VAR's for orders k and $k+1$.

A lag length of zero would imply that each season time series is made up of a constant plus white noise and it is therefore stationary, besides precluding any out of the year correlation among seasons. That possibility was clearly rejected for all variables. Besides producing a high value of the statistic when testing against a VAR(1) model, the model of constants left very substantial autocorrelation in the innovations of the seasons equations. Beyond that, a short VAR representation was chosen for most variables [Table 2, column 5]. The choices in the table are all consistent with lack of autocorrelation and partial cross correlations in the implied residuals.

The low orders that emerge from the likelihood ratio test are in conflict with the univariate representations of seasonality in section 2.2[6]. These conflicting results seem to indicate that, either we take as a fact the univariate representations, then implying that the use of the standard likelihood ratio test systematically leads to the wrong choice of VAR order, or seasonality has in fact a multivariate nature of possibly low order, a univariate representation then being inappropriate.

To avoid making a decision that might condition our results, we investigate both possibilities, the shorter VAR model for the seasons that emerges from the likelihood ratio test, and the longer VAR that is consistent with the previously estimated univariate representations.

3. TESTING FOR A UNIVARIATE REPRESENTATION OF SEASONALITY.

Once we have a VAR representation for the seasons, we propose a Likelihood Ratio Test of the joint set of restrictions on the covariance matrix and the VAR coefficients of the s-variate representation imposed by a univariate model of seasonality. We also propose a test of the subset of restrictions on the covariance matrix. Comparing the results of the two tests, we can draw some inference on whether a hypothetical rejection of the univariate representation is just due to heteroskedasticity among the seasons, or it has its foundation on the different dynamics of the seasons. We should bear in mind that rejecting the univariate ARIMA representation would still allow for the possibility that a periodic process as in Franses and Romijn (1993) could be accepted as the data generating process.

3.1 Homoskedasticity among the seasons.

It is well known that if $X_t = (x_{t1} \dots x_{ts})$ accepts a multivariate Gaussian VAR(p) representation, least squares yields consistent estimates even if X_t contains integrated and cointegrated variables. Moreover, in the presence of cointegration among $x_{t1}, \dots, x_{ts}$, the unconstrained least squares estimator has the same asymptotic distribution as the maximum likelihood estimator which observes the cointegration restriction [Park and Phillips (1989), Sims, Stock and Watson (1990), Lütkepohl (1993), p.369]. The s equations in each VAR correspond to the four quarters of a given variable and they are cointegrated in all cases, so these estimation results apply.

Under Normality, the independence of the ϵ -residuals introduced in (7) guarantees that the standard Bartlett[7] test for equal variances across samples, used here to test for equal variances across seasons, can be interpreted as a likelihood ratio test, which asymptotically follows a χ^2_{s-1} distribution. Rejection of the null hypothesis would imply that the diagonalized covariance matrix Λ is not scalar and hence, the analysis of the seasonal variable should be carried out in a multivariate framework.

Orthogonalizing the residuals as in (7) is crucial because, if the data generating process was indeed univariate, the resulting error term ϵ should be white noise. Lack of residual autocorrelation in the s-variate VAR(p) still allows for correlation among the seasonal residuals inside the same year. That would preclude their interpretation as being the innovations in a univariate model.

It is also important to realize that we are not searching for the univariate models of the seasons implied by the VAR model, and that standard procedures along that line would not allow us to carry on a test of the restrictions imposed by a univariate representation of seasonality. The interested reader should check appendix 1 for a more detailed discussion of this issue.

Columns 3 and 4 in Table 2 show the order of the VAR consistent with the estimated univariate models, and the value of the Bartlett statistic to test for equal variances

among seasons. At the 95% confidence level we reject the null hypothesis of variance homogeneity for all variables except: trade balance, exports, imports, vacancies, retail prices and the Financial Times index.

The results obtained with the shorter VAR models are shown in columns 5 and 6. We get now much more evidence of variance homogeneity, failing to reject the null hypothesis for the previous list, together with: consumption of durable goods, private investment, employment, average earnings and M4. For the remaining 12 variables there seems to be enough evidence of heteroskedasticity among quarters, consistent across VAR orders, which might be hard to reconcile with univariate representations of seasonality.

[INTRODUCE TABLE 2 HERE]

3.2 Homogeneity of the Seasons Equations and Homoskedasticity.

For a univariate model to be appropriate we need homoskedasticity across the seasons, but also that the statistical behavior of all seasons be the same. We now propose a general likelihood ratio statistic that jointly tests for both sets of constraints, i.e. homoskedasticity in the ϵ residuals as well as the validity of the dynamic cross- equation constraints that a univariate AR model representation imposes on a VAR for the seasons. Under Normality, we will have the *restricted log-likelihood function*:

$$\ln L_R = -\frac{ns}{2}\ln 2\pi - \frac{ns}{2}\ln \sigma_u^2 - \frac{1}{2\sigma_u^2}\sum_{j=1}^{ns} u_j^2$$

where u_j denotes the innovation in a univariate AR model.

If, on the contrary, a multivariate representation is needed then, again assuming Normality of the s-vector of innovations, we will have the *unrestricted log-likelihood*:

$$\ln L_U = -\frac{ns}{2}\ln 2\pi - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{n} a_t'\Sigma^{-1} a_t$$

where Σ is the s×s covariance matrix of the seasons innovations, and a_t denotes the vector of zero-mean residuals for the s seasons in year t. The value of the likelihood function is invariant to the linear transformation that changes the a_t into the ϵ_t residuals, so we can write it in terms of either the original (as above) or the orthogonalized VAR vector of innovations.

The maximum likelihood estimate (MLE) of the variance of u_j is:

$$\hat{\sigma}_u^2 = \frac{\sum_1^{ns} \hat{u}_j^2}{ns}$$

whereas the MLE of the covariance matrix Σ is:

$$\hat{\Sigma} = \frac{\sum_1^{n} \hat{a}_t \hat{a}_t'}{n}$$

which lead to the concentrated restricted and unrestricted likelihood functions:

$$\ln L_R = -\frac{ns}{2} \ln 2\pi - \frac{ns}{2} \ln \hat{\sigma}_u^2 - \frac{ns}{2}$$

$$\ln L_U = -\frac{ns}{2} \ln 2\pi - \frac{n}{2} \ln | \hat{\Sigma} | - \frac{ns}{2}$$

so that the likelihood ratio test for the univariate ARIMA seasonal model, as a restricted version of the multivariate VAR representation, consists on checking the significance of the statistic:

$$\lambda = -2 (\ln L_R - \ln L_U) = n \left[s \, . \ln \left(\frac{1}{ns} \sum_{j=1}^{ns} \hat{u}_j^2 \right) - \ln \left| \frac{1}{n} \sum_{t=1}^{n} \hat{a}_t \hat{a}_t' \right| \right]$$

This test is in the spirit of the test for time invariant coefficients and variances in a PAR model considered by Lütkepohl (1993), section 12.3.2. Ours is a similar test on an orthogonalized VAR model from which a PAR is a particular case. As in Lütkepohl, the likelihood ratio statistic λ above has a χ^2 asymptotic distribution with J degrees of freedom, J being the number of constraints, i.e., the difference between the number of estimated parameters in both models:

$$J = (s-1)(ps+2) + \frac{s(s-1)}{2} = (s-1)\left[s\left(p+\frac{1}{2}\right) + 2\right]$$

Table 3 shows the results of the test for a univariate nature of seasonality considering both VAR representations, the one with an order consistent with the estimated univariate model, as well as the one that results from choosing its order with a likelihood ratio test[8]. In addition, given the low number of effective observations, we take as a conservative criterium the use of a confidence level not higher than 90%.

[INTRODUCE TABLE 3 HERE]

With all these qualifications in mind, when the longer VAR models are used, eight variables seem to have a more complex seasonal pattern than can be captured by univariate models: total as well as private investment, value of physical increase in stocks, factor cost adjustment, personal disposable income, workforce, retail prices and the Financial Times index. When the alternative, lower order VAR models are used, we reject the univariate seasonal AR representation for all variables except Treasury bill interest rates.

At least in this context of a scarce number of observations, the order of the VAR representation used for the seasons is clearly a fundamental feature conditioning the results of the test. Higher order VAR representations tend not to reject the null hypothesis of a univariate nature of seasonality more often than the lower order ones. However, the results from the shorter VAR must be interpreted very carefully, since most univariate models in table 1 cannot be nested into them.

Nevertheless, the eight variables mentioned above emerge as having multivariate seasonality even when the longer VAR models are used as the alternative. From comparison with the results in Table 2, the rejection of univariate seasonality for six of them (total and private investment, physical increase of stocks, adjustment factor cost, personal disposable income and the workforce) is at least based on the presence of seasonal heteroskedasticity, while in the remaining two cases (retail prices and the Financial Times index), rejection seems to be based on season specific coefficients.

We mentioned in section 2.2 that two alternative univariate representations for each variable could be considered, a pure AR model and a mixed ARMA model, both in Osborn differences. One might prefer to replicate our analysis using mixed ARMA models, but as these models cannot be nested into VAR models for the seasons, estimation of VARMA models would be necessary. Appendix 1 shows, for the case of a VARMA(1,1) with two seasons, the implied constrains associated to a univariate mixed representation. Given the similar performance of both representations (pure AR and mixed univariate models) in fitting the data, we do not expect important qualitative differences with our results. We chose to compare pure AR univariate representations with VAR models to avoid the computational burden associated with specifying and estimating 23 VARMA processes.

The fact that the seasonal patterns in some variables may be more complex than can be captured by univariate representations, is consistent with recent work on testing for periodic structures in this same data set [Osborn et al. (1988), Osborn (1991), Franses and Romijn (1993) and Franses and Paap (1994)].

3.3 Analysis of extreme values.

Most variables in the sample contain large extreme values, that tend to produce evidence of heteroskedasticity among the seasons and hence, biasing our results towards misrejecting univariate seasonality. To check for the importance of their effects, we corrected the outliers beyond three standard deviations, using the intervention analysis of Box and Tiao (1975). Each input, multiplied by its estimated coefficient in the intervention model, was substracted from the original time series to obtain an outlier-corrected variable.

Heteroskedasticity results were sensible to the presence of outliers. As expected, correcting for outliers reduced the evidence of multivariate seasonality, with independence of the order of the VAR which was used in the joint test.

Under the longer VAR representations, six variables seemed to have multivariate seasonality after correction for outliers, versus the eight which were detected previously to the correction. With the shorter VAR models, 17 variables seemed to show multivariate seasonality, against the 22 that had that characteristic previous to outlier correction. Table 4 shows date and type of each intervention input.

[INTRODUCE TABLE 4 HERE]

Although extreme values tend to bias the results of the test towards multivariate seasonality, there is a subset of variables for which this feature is robust to the presence

of outliers. They are: total and private investment, physical increase of stocks, factor cost adjustment and personal disposable income.

4. CONCLUSIONS

A seasonal time series can be viewed as a realization coming from either a univariate or an s-variate stochastic process, s being the seasonal period.

The univariate approach leads to simple models that have been shown in the past to fit the data well. Nevertheless, this approach implies constant seasonal parameters and that may lead to substantial dynamic misspecification errors. On the other hand, although the multivariate approach allows for seasonally varying parameters, it leads to more complicated models that need longer time series to be elaborated.

Being concerned about the possibly multivariate nature of seasonality, we have proposed in this paper a procedure for testing the null hypothesis of a univariate versus an alternative multivariate representation of a seasonal time series. This is a joint test for the cross equation restrictions that a univariate model places on the coefficients of the s-variate representation, together with the constraints imposed on its contemporaneous error covariance matrix. It is however, quite simple to implement, since it becomes just a likelihood ratio test on the orthogonalized residuals from a VAR model for the seasons, compared with those from a univariate AR model. Combined with a test for homoskedasticity among seasons, it allows for further exploring whether the evidence of a multivariate nature of seasonality, when it arises, it is just due to a different variance for the seasons or it is caused by the dynamic correlations of the variable being season specific.

Studying 23 macroeconomic UK variables, we have found evidence against univariate seasonality in 8 cases (total as well as private investment, value of physical increase in stocks, factor cost adjustment, personal disposable income, workforce, retail prices and the Financial Times index), indicating that their seasonal patterns may be more complex than can be captured by univariate representations. This result is consistent with recent evidence on the presence of periodic structures in the same data set.

In implementing the test both, the choice of lag order for the VAR for the seasons as well as the correction of outliers, seem to be crucial in finding the proper character of seasonality. On one hand, standard likelihood ratio tests tend to select relatively low lag orders for the VAR processes which in turn lead to reject univariate representations often. On the other, the presence of outliers tends to bias the test towards multivariate seasonality.

NOTES

1. Similar results apply to more general VARMA processes, with minor changes. The simpler bivariate ARMA(1,1) case is discussed in Appendix 1.

2. Notice that the original single series x_t may be nonstationary. If so, the AR(8) polynomial will incorporate unit roots.

3. These polynomials will incorporate the appropriate constraints. For instance, a series with a unit root in the seasonal, as well as the non-seasonal frequencies, will need a $\Delta\Delta_4$ transformation to become stationary. That would imply a VAR seasonal representation of at least second order, since the double difference cannot be accomodated in a VAR(1).

4. For more details on the data, the reader can refer to Osborn (1990).

5. Only for vacancies we did not use the filter in Osborn (1990), who suggested this variable to be I(0). However, estimation of an univariate model under the assumption of stationarity failed to converge, due to the need of using a (1-B) filter. Therefore, we considered this variable to be I(1).

6. They are consistent, however, with recent research on this same sample [see Franses and Paap (1994)].

7. The Bartlett's statistic for the null hypothesis of equal variances:

$$H_0\colon \sigma_1^2 = \ldots = \sigma_s^2$$

is:

$$Q = (N-s)\ln S^2 - \sum_{i=1}^{s}(n_i-1)\ln S_i^2 = (n-1)\left[s\ln S^2 - \sum_{1}^{s}\ln S_i^2\right]$$

where S^2 is the sample variance computed with all observations, S_i^2 is the i-th season sample variance, n_i is the number of observations in each season, and N is the sample size (N=ns).

8. As in selecting the VAR lag order, in this case we have use the correction proposed in Sims(1980), i.e. the number **c** of estimated coefficients in each equation of the unrestricted VAR has been substracted from **n**, the number of years.

REFERENCES

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

Box, G.E.P. and G.M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, San Francisco, Holden Day.

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

Franses, P.H. and G. Romijn (1993), 'Periodic integration in quarterly UK macroeconomic variables', *International Journal of Forecasting*, 9, 467-476.

Franses, P.H. and R. Paap (1994), 'Model selection in periodic autoregressions' *Oxford Bulletin of Economics and Statistics*, 56, 4, 421-439.

Hillmer, S.E. and G.C. Tiao (1979), 'Likelihood function of stationary multiple autoregressive moving average models', *Journal of the American Statistical Association*, 74, 652-660.

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

Lütkepohl, H. (1993), *Introduction to multiple time series analysis*, Springer Verlag, Berlin.

Osborn, D.R., A.P.L. Chui, J.P. Smith and C.R. Birchenhall (1988), 'Seasonality and the order of integration for consumption', *Oxford Bulletin of Economics and Statistics*, 50, 361-377.

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

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

Park, J.Y. and P.C.B. Phillips (1989), 'Statistical inference in regressions with integrated processes: Part 2', *Econometric Theory*, 5, 95-131.

Sims, C.A. (1980), 'Macroeconomics and reality' *Econometrica*, 48, 1-49.

Sims, C.A, J.H. Stock and M. W. Watson (1990), 'Inference in time series models with some unit roots' *Econometrica*, 58, 113-144.

Theil, H. (1983), 'Linear algebra and matrix methods in econometrics', *Handbook of Econometrics*, Vol. I, 3-65.

Table 1
Pure seasonal AR representations on Osborn(1990) unit root filters

VARIABLE	Filter	Non-seasonal coefficients[b]				Seasonal coefficients[b]			σ_a%	Q(15)
		ϕ_1	ϕ_2	ϕ_3	ϕ_4	Φ_1	Φ_3	Φ_3		
GDP	$(1-B^4)$	.66 (.08)	.29 (.08)			-.61 (.08)	.41 (.08)		1.92	18.9
Comsumption	$(1-B^4)$	.81 (.09)	.17 (.09)			-.63 (.08)	-.29 (.09)		1.59	17.0
Durables	$(1-B^4)$	.66 (.09)	.18 (.09)			-.59 (.09)	-.31 (.08)		9.16	16.6
NonDurables	$(1-B^4)$	.98 (.02)				-.54 (.09)	-.21 (.09)		1.14	23.6
PubCons	$(1-B^2)$	.68 (.08)	-.72 (.08)	.68 (.09)	.24 (.09)	-.44 (.09)	-.27 (.09)		1.44	9.3
Investment	(1-B)	-.23 (.10)	-.16 (.10)			.52 (.09)	.27 (.09)		4.07	9.8
InvPriv	(1-B)	-.23 (.09)				.55 (.09)	.26 (.09)		5.25	15.6
InvPub	$(1-B^4)$	.66 (.10)	.20 (.10)			-.33 (.09)	-.40 (.09)		6.79	3.7
RSTB[a]	(1+B)	1.53 (.08)	-1.48 (.10)	1.34 (.11)	-.50 (.08)	-.50 (.09)			677.85	18.7
TradBal	(1-B)	-.36 (.08)				.23 (.08)	.25 (.08)		4.60	18.6
Exports	(1-B)	-.47 (.08)				.23 (.08)	.18 (.08)	.31 (.08)	3.99	16.7
Imports	(1-B)					.28 (.08)	.24 (.08)		4.13	16.0
Radjfc	(1-B)	-.40 (.08)	.26 (.08)			.62 (.08)	.23 (.08)		3.19	22.8
PDY	(1-B)	-.36 (.08)				.48 (.08)	.24 (.08)		2.18	28.4
Employment	(1-B)	.24 (.08)	.19 (.09)			.47 (.08)	.23 (.09)		0.50	21.1
Vacancies	(1-B)	.80 (.08)	-.08 (.11)	.12 (.11)	-.27 (.09)	.55 (.10)	.35 (.09)		8.38	17.0
WorkForce	(1-B)					.46 (.08)	.27 (.08)		0.42	18.6
Productivity	$(1-B^4)$	.63 (.08)	.31 (.08)			-.68 (.08)	-.42 (.08)		1.85	18.2
Avearn	(1-B)	.51 (.10)	.12 (.10)			.40 (.10)	.22 (.10)	.16 (.09)	1.27	16.7
RPI	$(1-B)^2$	-.56 (.08)	-.44 (.09)	-.43 (.09)	.34 (.09)	-.42 (.10)	-.16 (.09)		1.09	10.0
FTIndex	(1-B)	.24 (.10)							9.42	16.7
M4	$(1-B)^2$	-.66 (.09)	-.57 (.10)	-.54 (.10)	.31 (.09)	-.45 (.10)	-.23 (.10)		1.14	17.7
TBillyld[a]	(1-B)								1.36	21.7

[a] variables without the log-transformation
[b] standard errors in parentheses

Table 2

Tests for variance homogeneity among seasons

	n[b]	Lags p	Bartlett's test[c] for homoskedasticity	Lags p	Bartlett's test[c] for homoskedasticity
GDP	34	4	19.1	1	9.5
Consumption	34	4	44.6	1	12.6
Durables	34	4	15.5	1	2.9
NonDurables	34	4	10.1	1	12.4
PubCons	34	4	31.2	1	21.0
Investment	27	3	31.6	1	27.6
InvPriv	27	3	7.9	1	5.4
InvPub	27	4	52.0	2	17.7
RSTB[a]	34	3	18.2	2	15.3
TradBal	34	3	4.9	1	0.4
Exports	34	4	5.6	1	1.8
Imports	34	3	1.1	1	2.1
Radjfc	34	3	20.3	2	14.9
PDY	34	3	10.7	3	10.7
Employment	34	3	7.9	1	5.1
Vacancies	34	4	2.6	2	4.3
WorkForce	34	3	16.0	1	9.6
Productivity	34	4	11.0	1	8.0
Avearn	26	4	22.2	2	6.4
RPI	34	4	5.4	1	7.5
FTIndex	26	1	4.7	1	4.7
M4	26	4	14.7	1	3.4
TBillyld[a]	26	1	9.1	1	9.1

a variables without the log-transformation
b number of years in the sample
c the critical value at 95% confidence for a χ_3^2 is 7.8

Table 3

Tests for a univariate representation of seasonality

	n[b]	Lags p	Likelihood Ratio Statistic λ	Lags p	Likelihood Ratio Statistic λ
GDP	34	4	64.8	1	71.0
Consumption	34	4	62.1	1	51.2
Durables	34	4	65.3	1	51.6
NonDurables	34	4	65.0	1	49.8
PubCons	34	4	72.5	1	81.9
Investment	27	3	62.5	1	69.4
InvPriv	27	3	65.9	1	63.9
InvPub	27	4	51.7	2	78.0
RSTB[a]	34	3	80.9	2	88.0
TradBal	34	3	57.5	1	42.7
Exports	34	4	46.8	1	45.0
Imports	34	3	55.9	1	54.2
Radjfc	34	3	87.5	2	85.7
PDY	34	3	85.7	3	85.7
Employment	34	3	54.7	1	60.3
Vacancies	34	4	51.3	2	69.0
WorkForce	34	3	68.2	1	62.1
Productivity	34	4	64.6	1	69.7
Avearn	26	4	44.5	2	78.5
RPI	34	4	77.1	1	59.7
FTIndex	26	1	33.6	1	33.6
M4	26	4	25.7	1	38.9
TBillyld[a]	26	1	33.0	1	33.0

a variables without the log-transformation
b number of years in the sample
c the critical values at 90% confidence for a χ_J, J = 24, 36, 48 and 60 (associated to Lags = 1, 2, 3 and 4) are 33.1, 45.0, 62.0 and 74.4 respectively.

Table 4

Outliers

GDP	1/73 (S)
Consumption	2/68 (S); 1/73 (I); 1/78 (S); 2/79 (I)
Durables	2/68 (S); 2/79 (I)
NonDurables	2/68 (S); 1/73 (S); 1/78 (S); 2/80 (S)
PubCons	-
Investment	-
InvPriv	4/68 (I)
InvPub	1/73 (S); 2/87 (I)
RSTB	-
TradBal	4/66 (I); 4/67 (I); 1/79 (I)
Exports	4/67 (I); 2/69 (I); 3/72 (I); 1/79 (I)
Imports	4/66 (I); 3/81 (S)
Radjfc	1/78 (S); 2/80 (S)
PDY	2/72 (S); 2/76 (I)
Employment	2/59 (S)
Vacancies	4/66 (S); 1/74 (S); 1/75 (S); 2/75 (S); 1/80 (S)
WorkForce	2/59 (S); 1/71 (S); 3/83 (S)
Productivity	1/73 (S)
Avearn	1/74 (S); 2/74 (S); 3/74 (S); 4/74 (S)
RPI	1/75 (S); 2/75 (S); 3/79 (S)
FTIndex	1/75 (S); 2/75 (S); 1/77 (S); 4/87 (S)
M4	-
TBillyld	3/73 (S); 1/77 (S); 3/81 (S); 1/85 (S)

Note: S and I denote step and impulse type of interventions, respectively. Each one of these variables was affected by a single parameter in estimation.

APPENDIX 1

Autoregressive models

Let us assume, for simplicity, that we have just two seasons or semesters, i.e., two intra-year observations: x_{t1} and x_{t2}. If we have estimated a VAR(1):

$$\begin{pmatrix} 1-\phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1-\phi_{22}B \end{pmatrix} \begin{pmatrix} x_{t1} \\ x_{t2} \end{pmatrix} = \begin{pmatrix} a_{t1} \\ a_{t2} \end{pmatrix} \qquad \textbf{(A1)}$$

where B is the lag operator, and: $Var(a_{t1}) = \sigma_1^2$, $Var(a_{t2}) = \sigma_2^2$, $Cov(a_{t1},a_{t2}) = \sigma_{12}$, $\rho = \sigma_{12}/\sigma_1^2$, then, orthogonalizing with the Choleski decomposition would give us the system: where: $\epsilon_t = a_{t2} - \rho a_{t1}$. In this orthogonalized system, the first semester depends on the two

$$x_{t1} = \phi_{11}x_{t-1,1} + \phi_{12}x_{t-1,2} + a_{t1}$$

$$x_{t2} = \rho\, x_{t1} + (\phi_{21} - \rho\,\phi_{11})\, x_{t-1,1} + (\phi_{22} - \rho\,\phi_{12})\, x_{t-1,2} + \epsilon_t$$

previous observations of the complete time series, i.e., the two observations in the previous year. On the other hand, the second semester depends on the past three observations, i.e., the first semester of that same year, as well as the two observations from the previous year.

The observation previous to x_{t2} is x_{t1}, while that before x_{t1} is $x_{t-1\,2}$. Hence, this system of two equations collapses to a single equation, univariate model if and only if:

$$Var(\epsilon_t) = Var(a_{t1})$$

and the following cross equation restrictions hold:

$$\phi_{12} = \rho$$
$$\phi_{11} = \phi_{22} - \rho\,\phi_{12} \qquad \textbf{(A2)}$$
$$\phi_{21} = \rho\,\phi_{11}$$

In that case, we would have a univariate AR(2) model, with coefficients ϕ_{12} and ϕ_{11}. For a quarterly variable, this argument generalizes to produce an univariate AR(4) model out of a VAR(1) for the quarters, as we saw in section 2.

The standard procedure to derive univariate representations out of a multivariate time series model:

$$\Phi(B)\ X_t = a_t$$

is to rewrite it as:

$$D(B)\ X_t = [A(B)]^a\ a_t = \begin{pmatrix} 1-\phi_{22}B & \phi_{12}B \\ \phi_{21}B & 1-\phi_{11}B \end{pmatrix} \begin{pmatrix} a_{t1} \\ a_{t2} \end{pmatrix}$$

where D(B) denotes the determinant of the polynomial matrix $\Phi(B)$, and $[A(B)]^a$ is its adjoint matrix. In our example (A1), we get:

$$\begin{aligned} D(B)\ x_{t,1} &= a_{t,1} - \phi_{22}\ a_{t-1,1} + \phi_{12}\ a_{t-1,2} \\ D(B)\ x_{t,2} &= \phi_{21}\ a_{t-1,1} + a_{t,2} - \phi_{11}\ a_{t-1,2} \end{aligned} \qquad \text{(A3)}$$

where: $D(B) = (1-\phi_{11}B)(1-\phi_{22}B) - \phi_{21}\phi_{12}B^2$.

Equations (A3) give us univariate ARMA(2,1) models for x_{t1} and x_{t2}. However, a_{t1} and a_{t2} are contemporaneously correlated, so that the two cannot be interpreted as part of a same, univariate model, since its innovation would then necessarily be autocorrelated. To get rid of that autocorrelation, we would use the projection:

$$a_{t,2} = \rho\ a_{t,1} + \epsilon_{t,2}$$

with $\rho = \sigma_{12}/\sigma_1^2$, to rewrite (A3) as:

$$\begin{aligned} D(B)\ x_{t,1} &= \epsilon_{t,1} + \phi_{12}\epsilon_{t-1,2} + (\rho\phi_{12} - \phi_{22})\ \epsilon_{t-1,1} \\ D(B)\ x_{t,2} &= \epsilon_{t,2} + \rho\epsilon_{t,1} - \phi_{11}\epsilon_{t-1,2} + (\phi_{21} - \rho\phi_{11})\ \epsilon_{t-1,1} \end{aligned} \qquad \text{(A4)}$$

where $\epsilon_{t,1} = a_{t,1}$. System (A4) now reduces to a single univariate model for the whole time series **X** exactly under the same conditions (A2).

Note that the alternative projection: $a_{t1} = \beta a_{t2} + \epsilon_{t1}$ would not make much sense, since we would be projecting a_{t1} on a future observation, a_{t2}.

Mixed VARMA models

In the more general case of having estimated a mixed model, like a VARMA(1,1):

$$\begin{pmatrix} 1-\phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1-\phi_{22}B \end{pmatrix} \begin{pmatrix} x_{t1} \\ x_{t2} \end{pmatrix} = \begin{pmatrix} 1-\theta_{11}B & -\theta_{12}B \\ -\theta_{21}B & 1-\theta_{22}B \end{pmatrix} \begin{pmatrix} a_{t1} \\ a_{t2} \end{pmatrix}$$

with $Var(a_{t1}) = \sigma_1^2$, $Var(a_{t2}) = \sigma_2^2$, $Cov(a_{t1},a_{t2}) = \sigma_{12}$, $\rho = \sigma_{12}/\sigma_1^2$, which can be transformed into:

$$\begin{pmatrix} 1-\phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1-\phi_{22}B \end{pmatrix} \begin{pmatrix} x_{t1} \\ x_{t2} \end{pmatrix} = \begin{pmatrix} 1-\theta_{11}B & -\theta_{12}B \\ -\theta_{21}B & 1-\theta_{22}B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \rho & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} a_{t1} \\ a_{t2} \end{pmatrix} =$$

$$= \begin{pmatrix} 1-\theta_{11}B-\rho\ \theta_{12}B & -\theta_{12}B \\ -\theta_{21}B+\rho(1-\theta_{22}B) & 1-\theta_{22}B \end{pmatrix} \begin{pmatrix} \epsilon_{t1} \\ \epsilon_{t2} \end{pmatrix} = \begin{pmatrix} 1-(\theta_{11}+\rho\theta_{12})B & -\theta_{12}B \\ \rho-(\theta_{21}+\rho\theta_{22})B & 1-\theta_{22}B \end{pmatrix} \begin{pmatrix} \epsilon_{t1} \\ \epsilon_{t2} \end{pmatrix}$$

where $\epsilon_{t1} = a_{t1}$, $\epsilon_{t2} = a_{t2} - \rho a_{t1}$, which implies: $Cov(\epsilon_{t1},\epsilon_{t2}) = 0$.

In this representation with orthogonalized residuals, we would have:

$$\begin{aligned} x_{t1} &= \phi_{11}x_{t-1,1} + \phi_{12}x_{t-1,2} + \epsilon_{t1} - \theta_{12}\epsilon_{t-1,2} - (\theta_{11}+\rho\theta_{12})\ \epsilon_{t-1,1} \\ x_{t2} &= \phi_{21}x_{t-1,1} + \phi_{22}x_{t-1,2} + \epsilon_{t2} + \rho\epsilon_{t1} - \theta_{22}\epsilon_{t-1,2} - (\rho\theta_{22}+\theta_{21})\ \epsilon_{t-1,1} \end{aligned} \qquad \text{(A5)}$$

For these to collapse to a single equation, univariate mixed ARMA model we need some restrictions on the AR coefficients:

$$\begin{aligned} \phi_{11} &= \phi_{22} \\ \phi_{12} &= 0 \\ \phi_{21} &= 0 \end{aligned}$$

as well as on the MA parameters:

$$\begin{aligned} \rho &= -\theta_{12} \\ \theta_{22} &= \theta_{11} + \rho\theta_{12} \\ \rho\theta_{22} &= -\theta_{21} \end{aligned}$$

leading to a restricted model:

$$\begin{pmatrix} 1-\phi B & 0 \\ 0 & 1-\phi B \end{pmatrix} \begin{pmatrix} x_{t1} \\ x_{t2} \end{pmatrix} = \begin{pmatrix} 1-\theta_{22}B & \rho B \\ \rho & 1-\theta_{22}B \end{pmatrix} \begin{pmatrix} \epsilon_{t1} \\ \epsilon_{t2} \end{pmatrix}$$

In addition, the two error terms in (A5):

$$u_{t1} = \epsilon_{t1} + \rho\epsilon_{t-1,2} - \theta_{22}\epsilon_{t-1,1}$$
$$u_{t2} = \epsilon_{t2} + \rho\epsilon_{t,1} - \theta_{22}\epsilon_{t-1,2} \tag{A6}$$

need to have the same variance. That will happen if:

$$\mathrm{Var}(\epsilon_{t1}) = \mathrm{Var}(\epsilon_{t2}) \tag{A7}$$

as should be the case for the model to be univariate.

APPENDIX 2

Estimated univariate structures

Table A2.1 present mixed ARMA models estimated again under the differences in Osborn (1990). The mixed representations do not greatly reduce the number of coefficients and are very similar to the pure AR models in fitting the data. Residual standard deviations and Ljung-Box Q-statistics take similar values in both cases.

Comparing one or the other to their multivariate analogue should not be expected to produce any significant difference in our qualitative results.

Table A2.1
ARMA representations on Osborn(1990) unit root filters

		AR coefficients[b]						MA coefficients[b]				
		Non-seasonal				Seasonal		Non-seas.	Seasonal			
VARIABLE	Filter	ϕ_1	ϕ_2	ϕ_3	ϕ_4	Φ_1	Φ_2	θ_1	Θ_1	Θ_2	σ_a%	Q(15)
GDP	$(1-B^4)$	.99 (.01)						.36 (.08)	.72 (.06)		1.88	11.9
Consumption	$(1-B^4)$	.99 (.01)						.20 (.09)	.68 (.07)		1.57	11.7
Durables	$(1-B^4)$	.95 (.04)						.36 (.09)	.66 (.07)		9.04	14.8
NonDurables	$(1-B^4)$	.98 (.02)							.58 (.07)		1.12	20.0
PubCons	$(1-B^2)$	.96 (.03)	-.98 (.03)	.97 (.03)				.25 (.09)	.58 (.08)		1.43	7.5
Investment	(1-B)					.99 (.02)		.25 (.09)	.58 (.09)	.17 (.09)	3.88	4.7
InvPriv	(1-B)					.99 (.02)		.28 (.09)	.65 (.10)	.13 (.10)	4.84	10.7
InvPub	$(1-B^4)$	.90 (.06)				-.33 (.09)	-.41 (.09)	.22 (.12)			6.80	4.3
RSTB[a]	(1+B)	.53 (.07)				.26 (.09)	.30 (.09)	-.99 (.01)			677.8	18.0
TradBal	(1-B)					.94 (.05)		.42 (.08)	.77 (.10)		4.50	11.3
Exports	(1-B)					.98 (.03)		.52 (.07)	.82 (.07)		3.92	9.5
Imports	(1-B)					.96 (.04)		.19 (.08)	.79 (.08)		4.03	12.8
Radjfc	(1-B)					.97 (.03)		.48 (.08)	.41 (.09)	.20 (.09)	3.10	17.4
PDY	(1-B)					1.25 (.18)	-.29 (.15)	.35 (.08)	.79 (.14)		2.17	29.4
Employment	(1-B)	.28 (.08)	.24 (.08)			.99 (.01)			.81 (.06)		0.47	14.6
Vacancies	(1-B)	.69 (.06)				.97 (.02)			.64 (.08)		8.36	22.9
WorkForce	(1-B)					1.19 (.12)	-.19 (.11)		.84 (.07)		0.40	14.0
Productivity	$(1-B^4)$	.99 (.01)						.45 (.08)	.73 (.06)		1.82	14.7
Avearn	(1-B)	.57 (.08)				.94 (.04)			.62 (.10)		1.25	15.3
RPI	$(1-B)^2$					.98 (.03)		.59 (.07)	.85 (.07)		1.08	12.7
FTindex	(1-B)	.24 (.10)									9.42	16.7
M4	$(1-B)^2$	-.67 (.09)	-.61 (.10)	-.58 (.10)	.31 (.09)				.52 (.11)		1.14	18.2
TBillyld[a]	(1-B)										1.36	21.7

[a] variables without the log-transformation
[b] standard errors in parentheses

Abbreviations for variables in the tables:

RGDP: Gross Domestic Product at 1985 prices.

Consumption: Total personal expenditure on goods and services at 1985 prices.
Durables: Personal expenditure on durable goods at 1985 prices.
NonDurables: Personal expenditure on non-durable goods and services, defined as the total less expenditure on durables, both valued at 1985 prices.

PubCons: Total Government final consumption at 1985 prices.

Investment: Total gross fixed capital formation at 1985 prices.
InvPriv: Gross fixed capital formation of the private sector at 1985 prices.
InvPub: Gross fixed capital formation of the public sector, obtained as the sum of the series for general government and for public corporations, both at 1985 prices.

RSTB: Value of physical increase in stocks and work in progress at 1985 prices.

TradBal: Exports divided by imports of goods and services, both at 1985 prices.
Exports: Exports of goods and services, at 1985 prices.
Imports: Imports of goods and services, at 1985 prices.

Radjfc: Factor cost adjustment (taxes on expenditures less subsidies) at 1985 prices.

RPDY: Real personal disposable income at 1985 prices.

Employment: Workforce in employment, consisting of employees in employment, the self-employed, HM forces and participants in work-related government training programs.

Vacancies: Vacancies notified to job centres and remaining unfilled.

WorkForce: Workforce, consisting of the workforce in employment and the unemployed.

Productivity: GDP at 1985 prices divided by employment.

Avearn: Index of average earnings for the whole economy (1985=100).

RPI: Index of retail prices, all items (1985=100).

FTIndex: Financial Times index of industrial ordinary shares, (1 July 1935=100).

M4: Stock of 'broad' money, consisting of notes and coins in circulation plus private sector bank and building society deposits.

TBillyld: Percentage yield on Treasury bills.