Testing for time dependence in parameters

Ralf Becker

School of Economics and Finance

Queensland University of Technology

r.becker@qut.edu.au

Walter Enders

Department of Economics, Finance and Legal Studies

University of Alabama

wenders@cba.ua.edu

A. Stan Hurn

School of Economics and Finance

Queensland University of Technology

s.hurn@qut.edu.au

Abstract

This paper proposes a new test based on a Fourier series expansion to approximate the unknown functional form of a nonlinear time-series model. The test specifically allows for structural breaks, seasonal parameters and time-varying parameters. The test is shown to have very good size and power properties. However, it is not especially good in detecting nonlinearity in variables. As such, the test can help determine whether an observed rejection of the joint null hypothesis of linearity and time invariant parameters is due to time-varying coefficients or a nonlinearity in variables.

Keywords: time varying parameters. Fourier-series, nuisance parameters

JEL Classification: C51, C52 G12

1 Introduction

It is now generally agreed that linear econometric models do not capture the dynamic relationships present in many economic time-series. Moreover, adopting an incorrect non-linear specification may be more problematic than simply ignoring the non-linear structure in the data. It is not surprising, therefore, that non-linear model selection is an important area of current research and a wide array of testing procedures have been developed and subjected to rigorous empirical scrutiny (see, for example, Lee *et al.*, [1993] and Barnett *et al.,* [1997]). Although different in detail, these tests generally have two common attributes. *First,* the null hypothesis is taken to be a linear model with time-invariant parameters and is in fact, therefore, a joint hypothesis. *Second*, the available tests have good power against a wide variety of common nonlinear time-series models, but their ability to detect time dependence in the parameters, sometimes referred to as non-linearity in parameters, is, as yet, unknown. Given the vast number of plausible nonlinear specifications, to pare down all the potential nonlinear specifications to the actual nonlinearity present in the data generating process remains a difficult task (Ashley and Patterson, [2000]).

This paper proposes a new test based on a Fourier series expansion to approximate the unknown functional form of a time-varying autoregressive coefficient which seeks to provide a reliable test for time dependence in parameters. The null hypothesis remains a linear model with time-invariant parameters, while the alternative hypothesis specifies a model which is linear in variables but with time-varying coefficients. This specification contributes to the literature a number of important ways. In the first instance, time dependence in parameters is an important issue in its own right and includes issues related to structural breaks (Perron, [1989], Clements and Hendry, [1999]), seasonal parameters (Herwartz, [1995]) and time-varying parameters (Harvey, [1989]). The test proposed is capable of encompassing all these varying-parameter models in an integrated environment. A further important feature of this test is that, although it has good power against the alternative hypothesis of time dependence in parameters, it does not detect many other, important, nonlinear models. As such the test should help determine whether an observed rejection of the joint null hypothesis of linearity and time invariant parameters is due to time-varying coefficients or a nonlinearity in variables.

The issue of time-varying coefficients has been the subject of recent work by Lütkepohl and Herwartz [1996]. They propose a generalized flexible least squares (GFLS) estimation procedure, based on a standard likelihood function, which allows for time-varying coefficients by adding a series of penalty terms. By adjusting these penalties they are able to examine how coefficient variation contributes to the residual sum of squares. As Lütkepohl and Herwartz [1996] argue, however, GFLS is very much a tool for preliminary data analysis and visualisation; the method cannot provide a framework for statistical testing. Our use of a model with Fourier coefficients presents a method which allows identical conclusions to the GFLS method to be drawn in a statistically rigorous framework. Rigorous implementation of this testing strategy has to deal with the presence of unidentified parameters under the null distribution (Davies, [1987], Andrews and Ploberger, [1994], Hansen [1996]) and can be difficult to implement. It is demonstrated, however, that the problem may be restricted in such a way that a set of tabulated critical values can be applied without a meaningful loss of precision.

The paper is structured as follows. Section 2 outlines the GFLS method of Lütkepohl and Herwartz [1996]. Section 3 introduces the trigonometric testing approach and relates it to the literature of testing in the presence of unidentified parameters under the null and Section 4 outlines a simple implementation of the test which appears to yield satisfactory empirical results. In Section 5 a Monte Carlo exercise is performed to document the performance of the proposed procedure and an empirical example using the same data as Lütkepohl and Herwartz [1996] is presented in Section 6. Section 7 contains brief concluding comments.

2 Generalised Flexible Least Squares (GFLS)

Lütkepohl and Herwartz [1996] estimate $\boldsymbol{\alpha}_t$ in $y_t = \mathbf{y}_{t-1}\boldsymbol{\alpha}_t + \varepsilon_t$ by minimising the residual sum of squares plus two penalty terms, which penalise period to period $(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})$ and season to season $(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-S})$ changes in the parameters[1]. The objective function to be minimised in the parameter estimation process is

$$Q(\lambda_1, \lambda_2) = \sum_{t=1}^{T} (y_t - \mathbf{y}_{t-1}\boldsymbol{\alpha}_t)^2 + \lambda_1 \sum_{t=2}^{T} (\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})' \mathbf{D} (\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})$$
$$+ \lambda_2 \sum_{t=S+1}^{T} (\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-S})' \mathbf{D} (\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-S}).$$

where λ_1 and λ_2 are non-negative, fixed penalty parameters and $\mathbf{D}$ is a $(k \times k)$ weighting matrix. It is convenient to choose $\mathbf{D}$ to be a diagonal $(k \times k)$ matrix with diagonal elements

[1] Most of this paper will solely focus on time series applications, but the proposed testing and modelling strategies are readily applicable to any regression relationship.

equal to $\mathbf{y}'_{t-1,i}\mathbf{y}_{t-1,i}/T$ for $i = 1 \ldots k$. Further smoothness constraints can be imposed by adding longer lags in the penalty functions.

Estimation of the time path $\boldsymbol{\alpha}_t$, which minimises $Q(\lambda_1, \lambda_2)$, is achieved by means of a recursive algorithm, where values for the weighting matrix $\mathbf{D}$ and the penalty parameters are taken as given. Once the $(Tk \times 1)$ minimising parameter vector $\hat{\boldsymbol{\alpha}} = (\hat{\boldsymbol{\alpha}}_1, \hat{\boldsymbol{\alpha}}_2, \ldots, \hat{\boldsymbol{\alpha}}_T)$ is found, residual sum of squares, $RSS(\lambda_1, \lambda_2)$, can be computed. When λ_1 and λ_2 are very high and any parameter variation is prohibitively penalised, the resulting parameter estimates will be equivalent to OLS parameter estimates. When the penalties are inactive, $\lambda_1 = \lambda_2 = 0$, the parameters will vary such that the fit of the model to the data is perfect and $RSS(0, 0) = 0$.

The use of this method to detect time-variance in the parameters and, in particular, whether or not the parameters have a seasonal pattern, is best explained by using the simulation example reported in Lütkepohl and Herwartz [1996], p.274. Consider the following two models[2]:

Model 1: $AR[4]$

$$y_t = 0.9 y_{t-4} + \varepsilon_t, \qquad \sigma_\varepsilon^2 = 0.04\ . \tag{1}$$

[2] Here and in the following an AR[p] model is an autoregressive model which contains lag p only. This is in contrast to a AR(p) model, which includes all lags from 1 to p.

Model 2: Periodic Autoregressive(1) or PAR(1)

$$y_t = \alpha_{0s} + \alpha_{1s} y_{t-1} + \varepsilon_t, \qquad \sigma_\varepsilon^2 = 0.04 \tag{2}$$

with $\alpha_{0s} = -0.6, 0.3, -0.9, 0.8$

and $\alpha_{1s} = -0.4, 0.7, -0.3, 0.2$ for $s = 1, 2, 3, 4$.

Data are generated from these models ($T = 100$) and GFLS used to estimate $RSS(\lambda_1, \lambda_2)$ for different permutations of values for λ_1 and λ_2.

AR[4] Model

		λ_2			
		0	**0.001**	**1**	**1000**
	0		0	12.26	85.22
λ_1	**0.001**	0	0	12.56	85.26
	1	17.34	17.38	41.53	96.57
	1000	98.00	98.00	98.01	100.00

PAR(1) Model

		λ_2			
		0	**0.001**	**1**	**1000**
	0		0	3.07	13.77
λ_1	**0.001**	0	0	3.10	13.81
	1	25.90	25.91	32.79	49.95
	1000	98.96	98.96	98.97	100.00

Table 1: Normalised RSS obtained with objective function $Q(\lambda_1, \lambda_2)$ for $\lambda_1, \lambda_2 =$ 0, 0.001, 1, 1000 respectively. Normalisation enforces $RSS(1000, 1000) = 100$.

The flavour of these results is best obtained by focussing attention on the last column and the last row of the matrix of RSS values provided in Table 1 for each model. For the $AR[4]$ model, the imposition of the additional penalty on $(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-S})$, when $(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})$ is already heavily penalised (that is reading across the last row), does not result in a significant deterioration in the goodness of fit. Similarly, increasing λ_1 when $\lambda_2 = 1000$ (reading down the last column) has no significant impact in terms of RSS. By contrast, the lower panel of Table 1 reveals that the additional imposition of high penalties for period-to-period variation $(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})$ is extremely costly. This indicates that the restriction of $\boldsymbol{\alpha}_t \approx \boldsymbol{\alpha}_{t-1}$ is not supported by the data.

The methodology is designed to identify time-variability in parameters and to detect whether or not this variation is seasonal. However, as recognised by Lütkepohl and Herwartz [1996], GFLS does not itself admit statistical inference. Instead, detecting the presence of time-varying coefficients relies primarily on inspection of the patterns emerging from the variation of RSS with changes in the penalty parameters λ_1 and λ_2. Moreover, the estimated parameter time paths resulting from an application of GFLS depend crucially upon the chosen values of the penalty parameters. Since there is no way to determine the actual values of λ_1 and λ_2, there is no way to obtain the *true* time paths of the parameters. Once a parameter is identified as having seasonal variation, the $\{y_t\}$ series may be estimated using a specific non-linear time-series model as in Herwartz, [1995].

3 Tests based on Fourier approximation

This section develops an alternative approach for identifying parameter variation in a framework that allows for statistical inference on the estimated parameters. The idea of using trigonometric variables in regression relationships is not particularly new. For example, Gallant and Souza [1991] estimate Fourier parameters in a factor demand system; Engle [1974] and Tan and Ashley [1999] use frequency bands to test for parameter instability.

In a similar vein, we use a Fourier series expansion to approximate the unknown functional form of a time-varying autoregressive coefficient. Consider the following system

$$\begin{aligned} y_t &= \alpha_t y_{t-1} + \varepsilon_t \\ \alpha_t &= \alpha_0 + \sum_{j=1}^{M} \left[\alpha_{1j} \sin\left(\frac{2f_j\pi t}{T}\right) + \alpha_{2j} \cos\left(\frac{2f_j\pi t}{T}\right)\right]. \end{aligned} \tag{3}$$

If M is sufficiently large, the unknown functional form of α_t will be well approximated. If the null hypothesis $\alpha_{1j} = \alpha_{2j} = 0$ cannot be rejected for all values of j, then $\alpha_t = \alpha_0$ and the autoregressive coefficient is constant. Thus, instead of positing a specific model of parameter variation, the specification problem is to find the appropriate frequencies to include in the model. However, the choice of the appropriate number of frequencies to include in the Fourier approximation is not straightforward. To overcome this hurdle, we use a device first proposed by Ludlow and Enders [2000]. Specifically, we limit out attention to only one frequency so that (3) may be written as

$$y_t = \left[\alpha_0 + \alpha_1 \sin\left(\frac{2f\pi t}{T}\right) + \alpha_2 \cos\left(\frac{2f\pi t}{T}\right)\right] y_{t-1} + \varepsilon_t \tag{4}$$

If the value of f was known, it would be possible to test the null hypothesis $\alpha_1 = \alpha_2 = 0$ by means of a standard F-test. Rejecting the null hypothesis would imply a time-varying autoregressive coefficient.

In all reasonable circumstances, the true value of f is unknown and must be estimated from the data. The difficulty is that f is unidentified under the null hypothesis, in that different values of f do not change the likelihood of the data under the null (see Rothenberg, [1971]). It is therefore desirable to consider a range of values for this parameter f. A discrete set $\mathbf{\Gamma}$ needs to be specified and the testing strategy will take all values of $f \in \mathbf{\Gamma}$ into consideration. One consequence of this is that standard asymptotic theory cannot be used to derive the correct distribution of the test statistic under the null hypothesis. Alternative techniques to obtain the distribution of a test statistic must therefore be used and these are now outlined briefly.

3.1 The distribution of the test statistic

In implementing the general testing procedure, it is desirable to construct statistics which have an asymptotic chi-squared distribution such as the likelihood ratio, Lagrange multiplier and Wald statistics. For the purposes of exposition, consider the G likelihood ratio test statistics of the hypothesis $\alpha_1 = \alpha_2 = 0$ for all $f \in \mathbf{\Gamma}$, in (4). These may be denoted

$$LR(f_i) = -2(l_T - l_T^{f_i}) \qquad i = 1 \ldots G$$

where l_T is the log-likelihood of the restricted/linear model and $l_T^{f_i}$ is the log-likelihood of the model which includes the trigonometric variables. In order to make a statistical decision on the significance of the trigonometric terms, the information contained in $LR(f_1)$, $LR(f_2)$, ..., $LR(f_G)\}$ has to be distilled into one test statistic, LR_ϕ, whose distribution under the null hypothesis is known; this is achieved by specifying a mapping $\phi: R^G \rightarrow R$.

Note the two pieces of relevant information which will impact on the distribution of LR_ϕ

under the null distribution. *First*, it is well known that in large samples the distribution of each individual test statistic, $LR(f_i)$, under the null hypothesis is $\chi^2(2)$. *Second*, it is likely that

$$\mathrm{E}\left[LR(f_i), LR(f_j)\right] \neq 0 \qquad i \neq j.$$

In other words, the off-diagonal elements of the covariance matrix of the calculated likelihood ratio statistics, $\mathbf{\Sigma}$, may be non-zero (Hansen, [1996], [1999]) and this covariance structure may be application specific.

Once the mapping, ϕ, has been defined, Andrews and Ploberger [1994] and Stinchcombe and White [1998] show that the test statistic, LR_ϕ, and its distribution under the null hypothesis, $\chi(\phi, \Gamma, \Sigma)$, are given by

$$LR_\phi = \phi\left(LR(f_1), LR(f_2), \ldots, LR(f_G)\right) \tag{5}$$

$$\chi(\phi, \Gamma, \Sigma) = \phi(\chi^2_{2,1}, \chi^2_{2,2}, \ldots, \chi^2_{2,G}), \tag{6}$$

where each $\chi^2_{2,i}$ is a $\chi^2(2)$ deviate and

$$\mathrm{E}\left[\chi^2_{2,i}, \chi^2_{2,j}\right] = \mathbf{\Sigma} \qquad i,j{=}1\ldots G. \tag{7}$$

A statistical decision on the significance of any given test statistic, LR_ϕ, requires critical values obtained from the distribution $\chi(\phi, \Gamma, \mathbf{\Sigma})$, which is a non-trivial task (Davies, [1987], Andrews and Ploberger, [1994]). We follow Hansen [1999] and use two methods for obtaining the critical values. The *first* method draws random realisations from this asymptotic distribution in the following way. Construct G deviates from a $\chi^2(2)$ distribution ensuring that these deviates have covariance matrix $\mathbf{\Sigma}$, which is the covariance of the sample likelihood ratio statistics. One draw from the asymptotic distribution, $\chi(\phi, \Gamma, \mathbf{\Sigma})$, is then obtained by applying the mapping, $\phi(\cdot)$, to these $\chi^2(2)$ deviates. It is clear that the asymptotic distribution of the respective test statistics under the null hypothesis may be approximated by J draws and the proportion of these J realisations, therefore, which exceed the calculated value is an estimate of the appropriate p-value, $\widehat{p}$, which under standard conditions is approximately normal distributed with standard deviation $\sqrt{\hat{p}(1-\hat{p})/J}$. The *second* method for determining p-values is from bootstrap realisations of the test statistic, LR_ϕ. In this particular application, under the null hypothesis the model is a linear autoregressive one and bootstrap generation is relatively straightforward.

So far the mapping ϕ was left unspecified and indeed (5) and (6) are valid for a general class of functions ϕ (see Stinchcombe and White, [1998]). In this work, three variations of ϕ will be used, namely, the sup-norm, unweighted average and an exponentially weighted average of the arguments. In terms of the sample likelihood ratio statistics, the resultant test statistic under the mappings are given by

$$LR_{\text{sup}} = \sup_{f\in\Gamma} LR(f) \tag{8}$$

$$LR_{ave} = \frac{1}{G}\sum_{f\in\mathbf{\Gamma}} LR(f) \tag{9}$$

$$LR_{\text{exp}} = \ln\left[\frac{1}{G}\sum_{f\in\mathbf{\Gamma}} \exp\left(\frac{LR(f)}{2}\right)\right]. \tag{10}$$

Note that the latter two test statistics are derived under an optimal local power argument, LR_{ave} being the most powerful test for local alternatives and $LR_{\exp}$ performing superior for more distant alternatives.

3.2 Implementation

Consider the regression equation

$$\begin{aligned} y_t = {} & \alpha_{01} y_{t-1} + \alpha_{02} y_{t-2} + \ldots + \alpha_{0k} y_{t-k} \\ & + \alpha_1 \sin\left(\frac{2f\pi t}{T}\right) y_{t-p} + \alpha_2 \cos\left(\frac{2f\pi t}{T}\right) y_{t-p} + \varepsilon_t \end{aligned} \tag{11}$$

where k lags of y_t enter as ordinary, time-invariant, AR terms but the trigonometric terms operate only on one chosen lagged value of y_t, namely y_{t-p} for $p \in [1, k]$. It is now convenient to simplify notation slightly. Let

$\mathbf{y}_{t-1}$ = a $(1 \times k)$ vector of lagged values of y_t;

$\beta_0 = [\alpha_{01}, \alpha_{02}, \ldots \alpha_{0k}]'$;

$\mathbf{h}(y_{t-p},\ f) = \left[\sin\left(\frac{2f\pi t}{T}\right) y_{t-p},\ \cos\left(\frac{2f\pi t}{T}\right) y_{t-p}\right]$; and

$\beta_1 = [\alpha_1, \alpha_2]'$.

It is now possible to write equation (11) as

$$y_t = \mathbf{y}_{t-1}\boldsymbol{\beta}_0 + \mathbf{h}(y_{t-p}, f)\boldsymbol{\beta}_1 + \varepsilon_t \tag{12}$$

with the null hypothesis of interest being $\boldsymbol{\beta}_1 = 0$. Two further items of notation are needed.

Let

$\boldsymbol{\beta} = [\boldsymbol{\beta}_0 \ \boldsymbol{\beta}_1]'$,

$\mathbf{y}_{t-1}(f) = (\mathbf{y}_{t-1}, \ \mathbf{h}(y_{t-p}, f))$, and

$\mathbf{h}(f)$ the $(T \times 2)$ matrix of stacked $\mathbf{h}(y_{t-p}, f)$ for all $t = 1, ..., T$.

The vector $\mathbf{y}_{t-1}(f)$ may now be interpreted as the t^{th} row of the matrix $\mathbf{Y}_{-1}(f)$ which is obtained by stacking all T observations, $\mathbf{y}_{t-1}(f)$. Since the estimation of (12) is straightforward under both the null and alternative hypotheses, the computation of the G likelihood ratios $LR(f_i)$ and LR_ϕ is a relatively simple task.

In order to obtain critical values by means of the approach proposed by Hansen [1996], [1999], it is now necessary to construct the $\chi^2(2)$ deviates with covariance matrix $\boldsymbol{\Sigma}$. In order to generate the required deviates, Hansen [1999] suggests using the well known fact that the average regression score is normally distributed under the null hypothesis. In essence, $\chi^2(2)$ deviates may be constructed from squares of simulated regression scores with appropriate standardisation. Clearly the regression score of (12) will have $k+2$ elements of which only the last two, pertaining to the trigonometric variables $\mathbf{h}(f)$, are of interest. References to "score" should now be interpreted as relating only to these elements. By using $\mathbf{h}(f)$ in the simulation of regression scores, the required covariance structure for the $\chi^2(2)$ deviates will be ensured.

To generate a random realisation of the score under the null hypothesis, let $\mathbf{u}$ be $(T \times 1)$

random vector drawn from a standard normal distribution, set

$$\hat{\mathbf{u}} = \left(\mathbf{I} - \mathbf{Y}_{-1}(f)(\mathbf{Y}'_{-1}(f)\mathbf{Y}_{-1}(f))^{-1}\mathbf{Y}'_{-1}(f)\right)\mathbf{u} \tag{13}$$

and define the normalising factor

$$\mathbf{M}(f)^{-1} = \left[\mathbf{h}(f)'\mathbf{h}(f) - (\mathbf{h}(f)'\mathbf{Y}_{-1})(\mathbf{Y}'_{-1}\mathbf{Y}_{-1})^{-1}\mathbf{h}(f)\right]^{-1}$$

as the lower right hand (2×2) submatrix of $(\mathbf{Y}_{-1}(f)'\mathbf{Y}_{-1}(f))^{-1}$ (see Hendry, [1995], for the rules of partitioned inversion).

A random realisation of the average score can be generate by means of $\hat{\mathbf{u}}'\mathbf{h}(f)/T$. Due to the projection in (13) the average score generated in this manner will be zero, as required under the null hypothesis. The suitably standardised squared version of this average score, $\hat{\mathbf{u}}'\mathbf{h}(f)\mathbf{M}(f)^{-1}\mathbf{h}(f)'\hat{\mathbf{u}}$, is the required $\chi^2(2)$ deviate. *One* draw from the asymptotic null distribution of LR_ϕ is then[3]

$$\phi\left\{\hat{\mathbf{u}}'\mathbf{h}(f_1)\mathbf{M}(f_1)^{-1}\mathbf{h}(f_1)'\hat{\mathbf{u}}, \ldots, \hat{\mathbf{u}}'\mathbf{h}(f_G)\mathbf{M}(f_G)^{-1}\mathbf{h}(f_G)'\hat{\mathbf{u}}\right\}, \tag{14}$$

and J realisations of this process will enable appropriate p-values to be computed.

The second method for determining p-values is from bootstrap realisations of the test statistic. Since the null model is a model of the autoregressive class, the bootstrap generation

[3] This is valid for a stationary model with homoscedastic residual terms. Detailed derivations of the tests null distribution are given in Hansen [1996], Andrews and Ploberger [1994] and Andrews [1994].

has to follow a recursive scheme, using appropriate starting values and assuming the validity of the null hypothesis $\boldsymbol{\beta}_1 = 0$.

$$y_t^* = \mathbf{y}_{t-1}^* \hat{\boldsymbol{\beta}}_0 + \varepsilon_t^*$$

The bootstrapped residuals ε_t^* are resampled (with replacement) from the empirical distribution of the $\hat{\varepsilon}_t$s. The p-value for the LR_ϕ test statistic computed from the original data is again estimated by the proportion of the J bootstrapped test statistics.

The consistency of this bootstrap method has been proven by Bose [1988], when applied to estimating the distribution of $\boldsymbol{\beta}_0$. It is well known that LR test statistics are asymptotically pivotal and bootstrap techniques therefore promise to deliver tests with good statistical properties (Li and Maddala, [1996]). Here, however, the test statistic is a mapping of correlated LR test statistics and to the best of our knowledge, there is no theoretical work establishing consistency of bootstrap methods applied to such test statistics. The approach in this paper is purely empirical, in that the correct size of the bootstrap test is established by means of a Monte Carlo simulation.

4 A simple implementation of the test

The previous section has described how the distribution of the test statistic may be obtained by asymptotic draws from the specific distribution under the null hypothesis or by bootstrapping. This testing procedure is technically correct, but it is difficult to implement using standard software packages. This section discusses an alternative way of implementing the test.

In one special case, established by Davis [1987], the distribution of the test statistic is invariant under the null hypothesis and hence it is possible to tabulate critical values. Although in other cases considered here the distribution of the test statistic under the null hypothesis is not invariant, it transpires that the variation in critical values is so slight that it is possible to use tabulated critical values for all practical purposes.

It is clear from equations (6) and (7) that an invariant distribution under the null hypothesis is obtained when $\Sigma = I_G$, or, in other words, $\mathrm{E}[LR(f_i), LR(f_j)] = 0$ for $i \neq j$. Consider the specification used by Davis [1987]

$$y_t = \alpha_0 + \alpha_1 \sin\left(\frac{2f\pi t}{T}\right) + \alpha_2 \cos\left(\frac{2f\pi t}{T}\right) + \varepsilon_t \tag{15}$$

where ε_t is independently normally distributed. If the domain of the frequency, f, is restricted to integer values in the interval $1 \ldots T/2$, then the trigonometric terms in equation (15) are all orthogonal at all points in this restricted domain of f. This result may be used to generate a test for $\alpha_1 = \alpha_2 = 0$ in the following way

Estimate equation (15) by OLS for each value of f in the restricted domain. Denote by f^* the frequency which yields the smallest residual sum of squares, RSS^*, and let α_0^*, α_1^*, and α_2^* be the coefficients associated with that frequency value. An F-type test, F_{trig} , of $\alpha_1 = \alpha_2 = 0$ is given by

$$F_{trig}(f^*) = \frac{(RSS_r - RSS^*)/2}{RSS^*/(T-k+1)}$$

where RSS is the residual sum of squares with the restriction imposed. Since the F_{trig}

statistic is calculated for the frequency which minimised the residual sum of squares, this is equivalent to using the sup norm mapping of the previous section. Note, however, that, although invariant, the distribution of the test statistic under the null hypothesis does not follow a standard F-distribution with 2 and $T-k+1$ degrees of freedom, and critical values need to be tabulated by simulation[4].

Unfortunately the special case of equation (15) simply tests the constant term for time dependence. The more interesting applications of this testing procedure, at least in this paper, are aimed at identifying time variation in autoregressive coefficients[5]. While it is still possible to implement the simple testing strategy based on an integer frequency chosen by minimising the residual sum of squares, the orthogonality property of the trigonometric terms does not carry over. It follows therefore that the distribution of the test statistic under the null hypothesis is no longer invariant. It transpires, however, that the variation in the critical values of the distribution of F_{trig} are so small that for all practical purposes tabulated critical values may be used without significant loss of precision.

The effect on the critical values of the F_{trig} test of changing the parameter values of the underlying linear model specified as the null hypothesis, is easily demonstrated as follows.

[4] Davies [1987] derives the invariant distribution under the null hypothesis for this special case based on a test statistic which is distributed as a χ_2^2 deviate for any fixed frequency.

[5] Of course the test may also be used to test any parameter in any regression model for time dependence.

	Critical Values for $\mathbf{F}_{trig}$ test			
	AR(1)	**AR(2)**	**AR(2)**	**AR[4]**
	$H_0 : \alpha_{1t} = \alpha_1$	$H_0 : \alpha_{1t} = \alpha_1$	$H_0 : \alpha_{2t} = \alpha_2$	$H_0 : \alpha_{4t} = \alpha_4$
Parameter				
range				
0.1	7.206	7.282	7.280	7.265
0.2	7.206	7.320	7.274	7.265
0.3	7.248	7.322	7.277	7.254
0.4	7.237	7.304	7.273	7.249
0.5	7.236	7.305	7.265	7.255
0.6	7.273	7.331	7.289	7.269
0.7	7.284	7.337	7.299	7.271
0.8	7.292	7.346	7.308	7.287
0.9	7.278	7.314	7.303	7.314

Table 2: 5autorgressive models and parameter values.

The following autoregressive models were simulated for the parameter ranges shown:

$$AR(1) : y_t = \alpha_0 + \alpha_1 y_{t-1} + \epsilon_{1t} \qquad \forall\, \alpha_1 \in [0.1, 0.9]\ ,$$

$$AR(2) : y_t = \alpha_0 + \alpha_1 y_{t-1} - 0.3 y_{t-1} + \epsilon_{1t} \qquad \forall\, \alpha_1 \in [0.1, 0.9]\ ,$$

$$AR[4] : y_t = \alpha_0 + \alpha_4 y_{t-4} + \epsilon_{4t} \qquad \forall\, \alpha_4 \in [0.1, 0.9]\ .$$

In each case the 5% critical values of the F_{trig} test was computed from 50000 repetitions. Note that the $AR(2)$ model allows the test to be conducted on both $\alpha_1$and α_2. The results are presented in Table 2.

It is clear from these results that changing the parameters of the underlying linear models results in only minor variations in the critical values of the distribution of the test statistic under the null hypothesis. It appears, therefore, that this simplified procedure is a viable practical approximation to the generation of the correct asymptotic distribution outlined in the previous section. The key to this result is the restriction of the domain of the frequency, f, to only integer values between 1 and $T/2$. The empirical performance of the F_{trig} test will be evaluated in the next Section along with the LR_{ϕ} tests. As the approximate distribution changes very little with changes in the parameters, the critical values tabulated in Ludlow and Enders [2000], where only variations the sample size for given Γ are considered, can be used.

5 Empirical properties of the test

This section reports results from Monte Carlo experiments which were designed to demonstrate the following points. *First*, the F_{trig} and the $LR_{\sup}$, LR_{ave} and $LR_{\exp}$ tests have the correct statistical size under the null hypothesis of a linear time-invariant autoregressive process. *Second*, the tests have significant power properties for the alternative hypothesis of time-varying coefficients. As a by-product of this inquiry we show that these tests also have power against alternative hypotheses which are non-linear in variables.

In order to provide results comparable to Lütkepohl and Herwartz [1996] the two artificial examples given in Section 2 of their paper are used again[6]. Recall that in process (1) both the parameters were time-invariant and in process (2) both parameters displayed seasonal patterns. The processes with sample size 100 were simulated 5000 times and for each realisation the constant and the autocorrelation coefficients were tested for time invariance

[6] They also provide a third example, which also incorporates a structural break. Dealing with structural breaks is an important aspect of time-series modelling but not objective of this paper.

process	Test	Draw from asymp. distribution			Bootstrap		
	F_{trig}	$LR_{\sup}$	LR_{ave}	LR_{exp}	$LR_{\sup}$	LR_{ave}	LR_{exp}
$AR(1)$							
0.01	0.0106	0.016	0.002	0.016	0.007	0.013	0.008
0.05	0.0572	0.074	0.036	0.070	0.058	0.057	0.052
0.10	0.1196	0.134	0.085	0.122	0.109	0.108	0.112
$AR[4]$							
0.01	0.011	0.013	0.002	0.014	0.008	0.008	0.008
0.05	0.053	0.065	0.030	0.065	0.044	0.048	0.051
0.10	0.104	0.13	0.076	0.118	0.100	0.093	0.098
$AR[4] - const$							
0.01	0.013	0.018	0.000	0.018	0.008	0.008	0.008
0.05	0.069	0.082	0.000	0.076	0.044	0.048	0.051
0.10	0.127	0.141	0.000	0.134	0.100	0.093	0.098
$PAR(1)$							
0.01	1.0	1.0	0.015	1.0	1.0	0.041	1.0
0.05	1.0	1.0	0.098	1.0	1.0	0.117	1.0
0.10	1.0	1.0	0.198	1.0	1.0	0.208	1.0
$PAR(1) - const$							
0.01	1.0	1.0	0.497	1.0	1.0	0.531	1.0
0.05	1.0	1.0	0.834	1.0	1.0	0.571	1.0
0.10	1.0	1.0	0.945	1.0	1.0	0.587	1.0

Table 3: Simulation results of the F_{trig}, $LR_{\sup}$, LR_{ave} and LR_{exp} tests applied to AR(1), AR[4] and PAR(1). Where not specified otherwise, the AR coefficient is tested for time-invariance.

in the described manner. Due to the increased computational burden, the results for the LR tests were based on 1000 simulations ($J = 500$).

The simulation results of the OLS based F-test and the LR tests described in the previous section are reported in Table 2, which also includes the simulation results for a conventional AR(1) process. The AR[4] process has time-invariant parameters and therefore is a representation of the null hypothesis. The rejection frequencies are close to the expected frequencies, indicating that the tests have approximately the correct size. The only exception is theLR_{ave} test applied to the constant of the $AR[4]$ model, which is far too conservative. If any general conclusion must be drawn, then it appears that the empirical size of the LR tests based on bootstrapped p-values is slightly superior to the other two tests. The time variance in the parameters of the $PAR(1)$ model is easily detected by all tests except the LR_{ave} test. This is to be expected, because this test is particularly powerful against local alternatives and the $PAR(1)$ model is the most global alternative possible. For the other tests, the null hypothesis of time invariance in both coefficients of the $PAR(1)$ process was rejected in 100% of the cases. More importantly, when considering quarterly seasonality in the coefficients, the optimal frequency f^* in all cases is 25. In conjunction with the sample size 100 this implies a period of $100/25 = 4$ for the coefficient variation. The coefficients therefore have a quarterly-varying component. The efficacy of the F_{trig} test is particularly pleasing given its OLS foundations.

In order to scrutinise the power properties of the test two further time-varying coefficient models were simulated

$$y_t = \alpha_t \ y_{t-1} + \varepsilon_t$$

$$\mathrm{SC} : \alpha_t = -0.2(0.9) \text{ for } t = 4, 8, 12 \ etc. (\text{all other } t)$$

$$\mathrm{ARC} : \alpha_t = 0.3 + 0.5\alpha_{t-1} + \nu_t, \ \nu_t \sim N(0, 0.25).$$

SC is a seasonal coefficient model and ARC features a coefficient following a stationary autoregressive model of order one. In the introduction it was mentioned that time-varying coefficient models and models nonlinear in variables are treated separately in the literature. And indeed its modelling principles are different. However, in real life situations it might be difficult to, per-se, rule out one or the other. In the following it will be demonstrated that, provided no firm theoretical grounding is available to dismiss one model type, it might be difficult to differentiate between them on the basis of testing alone. For that purpose we extend the simulation by three processes, which are nonlinear in variables. These are

$$\text{BL} : y_t = 0.7\ y_{t-1}\varepsilon_{t-2} + \varepsilon_t$$

$$\text{TAR} : y_t = 0.9\ y_{t-1} + \varepsilon_t$$

$$a_t = 0.9(-0.3\ )\ for\ |y_{t-1}| \square\ 1(> 1)$$

$$\text{LSTAR} : y_t = (0.0 + 0.02F_t) + (1.8 - 0.9F_t)\ y_{t-1}$$

$$+(-1.06 + 0.795F_t)\ y_{t-2} + \varepsilon_t$$

$$\text{where } F_t = [1 + \exp(100\,(y_{t-1} - 0.02))]^{-1}$$

$$\text{and } \varepsilon_t \sim N(0, 0.02^2)$$

and have been used in simulation studies by Lee *et al.* [1993], Teräsvirta *et al.* [1993] and Dahl [1998]. In Table 3 simulation results for the F_{trig}, and both versions of the $LR_{\exp}$ test are reported. In addition, commonly used tests of nonlinearity in variables, namely the

process	Test F_{trig}	LR_{exp}	LR^b_{exp}	V23	Tsay	LSTAR	CUSUM	$CUSUM^2$
SC								
0.01	0.932	0.944	0.925	0.227	0.023	0.880	0.002	0.131
0.05	0.971	0.979	0.967	0.399	0.091	0.959	0.016	0.283
0.10	0.985	0.987	0.980	0.491	0.158	0.979	0.040	0.393
ARC								
0.01	0.702	0.450	0.740	0.348	0.162	0.349	0.058	0.548
0.05	0.825	0.662	0.873	0.471	0.258	0.456	0.128	0.662
0.10	0.878	0.765	0.916	0.549	0.326	0.524	0.194	0.730
LSTAR								
0.01	0.042	0.243	0.198	0.854	0.856	0.880	0.011	0.085
0.05	0.129	0.440	0.378	0.942	0.943	0.959	0.055	0.210
0.10	0.209	0.541	0.502	0.967	0.970	0.979	0.116	0.303
TAR								
0.01	0.010	0.017	0.012	0.228	0.008	0.009	0.004	0.007
0.05	0.053	0.056	0.041	0.455	0.041	0.052	0.036	0.044
0.10	0.103	0.108	0.090	0.595	0.091	0.099	0.077	0.091
BILIN								
0.01	0.609	0.673	0.627	0.266	0.173	0.892	0.022	0.517
0.05	0.770	0.826	0.815	0.416	0.292	0.949	0.066	0.674
0.10	0.834	0.895	0.897	0.514	0.374	0.964	0.114	0.753

Table 4: Power statistics for trigonometric tests. The two versions of the LR test are labelled a and b respectively.

Tsay test, the V23 and the LSTAR test, are also applied for comparative purposes. All three tests[7] examine the significance of cross products of elements in $\mathbf{Y}_{-1}$. A widely used test for parameter instability is the CUSUM and the $CUSUM^2$ test (Brown *et al.,* [1975]).

The results reported in Table 3 confirm that the tests basing on trigonometric expansions, detect time-variance in parameters reliably. They are also far superior to the CUSUM and the $CUSUM^2$ tests. In fact, in the context of the time varying parameter processes simulated here, the CUSUM test can hardly be recommended. However, it should be noted that the trigonometric tests also have power against the bilinear model. This is not surprising, since

[7] A good overview for these tests is given in Teräsvirta et al [1993].

the latter can be interpreted as a model whose AR(1) coefficient is randomly drawn from a normal distribution with mean 0.7 and variance equal to the variance of ε_{t-1}. Therefore the coefficient is time varying, but in an unsystematic way. The tests power against the TAR and the LSTAR model is, as we would have hoped, very limited. This facilitates the differentiation between a time-varying coefficient model and a model nonlinear in variables.

From these results it can be concluded, that in real-life problems, it might be difficult to rule out one of the two process classes on the basis of testing results. Also it appears that F_{trig} test performs as well as the LR tests based on the asymptotic or bootstrap distribution, which are more difficult to implement and also more computing intensive. The F_{trig} test does not require more than the ability to calculate OLS regressions over a range of values for f and to store the obtained results. This is easy to implement in most standard econometric packages.

6 Empirical application

Lütkepohl and Herwartz [1996] illustrate their procedure using seasonally unadjusted quarterly per-capita non-entrepreneurial income and consumption data from1960.I to 1988.IV (116 observations). Although the time-period straddles German unification, we use the identical span so as to facilitate comparison of the two methods.[8]

[8] We thank Hiltrud Nehls from the Rhine-Westphalia Institute for Economic Research for providing the data. The data are in actual prices.

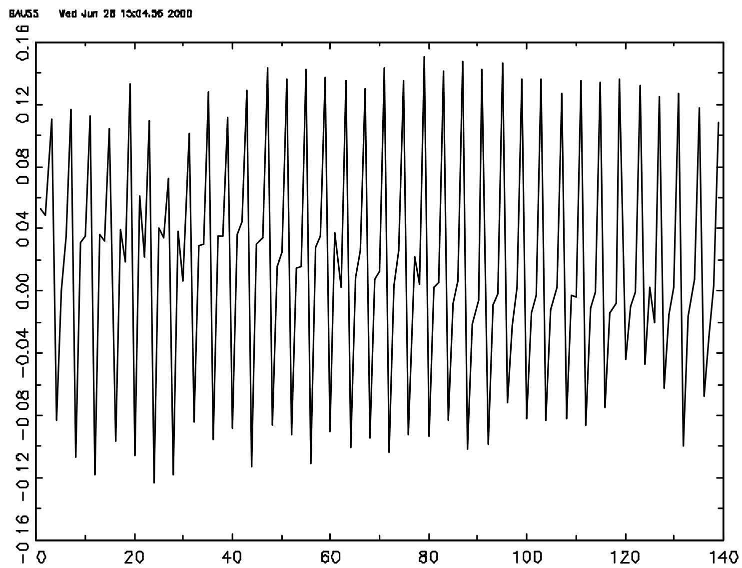

Figure 1: Figure 1: Difference in log per-capita income. 1960.I to 1994.IV.

6.1 German per-capita income

The log difference of the quarterly income data display an obvious quarterly pattern (Figure 1). Growth rates are negative for QI and in general positive for the remaining quarters.

We will start the testing procedure with an AR(1) model[9]. The null of time invariance is rejected for the constant and the AR(1) coefficient. The test statistics are 60.8 and 20.1 respectively[10]. The frequencies $f^*$are 28 and 29 respectively and, in combination with $T = 115$, hint at a quarterly variation. Lütkepohl and Herwartz use an AR(4) model as their base model and conclude that only the constant term displays seasonal variation. The

[9] In general only the results of the $F(f^*)$ test are reported. Test results basing on the bootstrap distribution are only reported when they indicate different decisions.

[10] The critical values for a sample size of 100 and a frequency search region of $k = 1 \ldots T/2$ are 6.37 (10% significance), 7.19 (5%) and 9.09 (1%) (Table 1 in Ludlow and Enders, [2000]).

$F(f^*)$ test applied to the AR(4) model does not reject the null hypothesis of parameter constancy for the four autoregressive coefficients. All test statistics are well below the 10% critical value. The test statistic for the constant is marginally significant at the 10% level. However, the selected frequency is $f^* = 35$, which is too high to indicate a seasonal behaviour. The LR bootstrap tests confirm this marginal result. While LR_{ave} is clearly nonsignificant, the $LR_{\sup}$test statistic has a p-value of 0.102 and the $LR_{\exp}$ test has an estimated p-value of 0.072.

If the goal was to further model the process driving the quarterly growth rate of German per-capita income, one would have to decide whether one should use the AR(4) model as the basis model, where the inclusion of the autoregressive component of order 4 appears to be sufficient to capture most of the seasonality. Alternatively, a periodic autoregressive model of order one might be the appropriate model. The choice is made difficult because it is well known (Herwartz, [1995] and Franses, [1996]), that PAR models create autocorrelation patterns which might well be approximated by a higher order linear AR model. One would have to resort to criteria such as forecast performance or economic plausibility to decide for one or the other model.

6.2 German per-capita consumption

From Figure 2 the seasonal pattern in the log difference in consumption, c, is obvious. Note that the pattern of the seasonal variation appears to be relatively stable.

When testing the AR(1) model for time constant parameters, the results are again clear cut. The null of time invariance is rejected for both coefficients with test statistics of 46.2

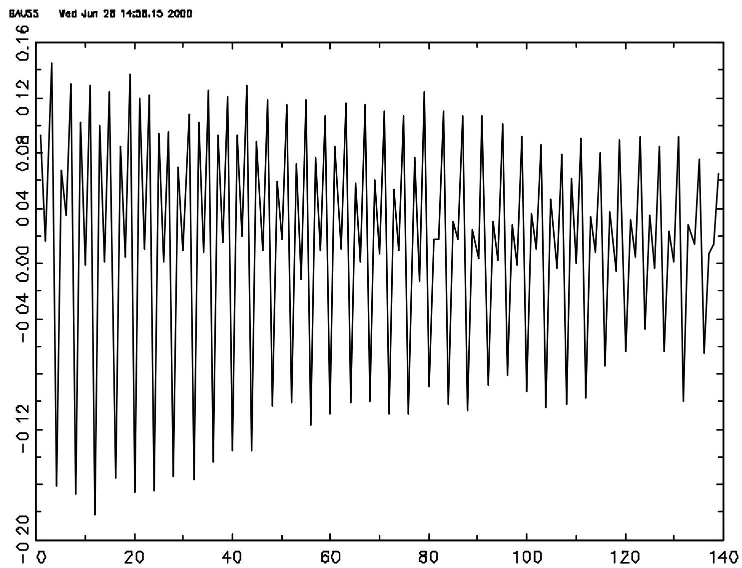

Figure 2: Figure 2: Difference in log per-capita consumption. 1960.I to 1994.IV.

parameters	$F(f^*)$	LR_{sup}	LR_{ave}	LR_{exp}	f^*
constant	9.3^{**}	0.008	0.102	0.01	45
α_1	15.8^{**}	0.000	0.014	0.000	10
α_2	13.3^{**}	0.000	0.046	0.000	10
α_3	6.5^{*}	0.146	0.912	0.178	10
α_4	10.8^{**}	0.010	0.486	0.010	10

Table 5: Test on parameter constancy in AR(4) model for per-capita consumption data. Test statistics for the F_{trig} test and p-values otherwise.

and 21.3 respectively. Again the frequencies f^* (28 and 29) clearly indicate a predominant seasonal variation. When testing the coefficients of an AR(4) model for time constancy, results are markedly different to the results in the income data.

The null of time constancy is clearly rejected for the constant and the coefficients α_1, α_2 and α_4. This result is slightly different Lütkepohl and Herwartz's conclusion, who after visually evaluating the fluctuations in parameters, conclude that α_1, α_2 can be treated as constant. The test results presented here do not support this. All tests, including the LR_{ave}

test[11] reject the null hypothesis for α_1 and α_2. Only the $F(f^*)$ test rejects $\alpha_3 = 0$ marginally. The frequencies f^*, however, do not support a quarterly variation in the coefficients. The AR(4) model has 111 usable observations. In conjunction with $f^* = 10$ this hints at cycles in the autoregressive parameters of approximately 3 years. Interestingly, when AIC or SIC criteria are used to find the optimal AR lag structure, both criteria select lag 12. This and the $f^*$are snapshots of the same phenomenon.

It is not in the scope of this paper to investigate this further and, for example, examine whether possibly a low order PAR model is suitable. It is known that the latter can create autocorrelation patterns at multiples of the number of seasons. The purpose of this exercise was to illustrate how the results obtained by means of GFLS can be achieved with much simpler means. Also, f^* which is a natural side product of the proposed testing strategy, carries valuable information.

7 Conclusion

The testing framework presented here provides a useful framework for testing regression parameters for time dependence. A trigonometric approximation is shown to capture the effects of possible variation in the parameters of a time-series model and allows the reliable detection of any such variation in a rigorous statistical manner. The primary econometric difficulty encountered in implementing this testing strategy is that the appropriate frequency in the Fourier approximation is unidentified under the null hypothesis. As a result the test

[11] All significance statements for the LR tests are based on the bootstrapped distribution under the null hypothesis.

statistic has a non-standard distribution under the null hypothesis. It is possible to assess the significance of the test statistic either by draws from the asymptotic distribution under the null hypothesis or by bootstrapping. Alternatively, if the domain of the frequency is limited to integer values it is shown that the critical values of the test statistic's distribution is relatively stable, at least for the models examined in this paper. As a result tabulated critical values constructed by Ludlow and Enders [2000] work well in the sense that the test's empirical performance is not impaired by this simple approximation.

The test is shown to have very good size and power properties. However, it is not especially good in detecting nonlinearity in variables. As such, the test can help determine whether an observed rejection of the joint null hypothesis of linearity and time invariant parameters is due to time-varying coefficients or a nonlinearity in variables. As an example, we used the same West German income and consumption data as Lütkepohl and Herwartz [1996]. It was demonstrated that the results generated by Generalised Flexible Least Squares may be obtained in a statistically rigorous manner.

References

[1994] Andrews D.W.K. (1994) "Empirical process methods in econometrics" in: Engle R.F. and McFadden D.L. *Handbook of Econometrics, Volume IV*, Elsevier Science: Amsterdam.

[1994] Andrews D.W.K. and Ploberger W. (1994) "Optimal tests when a nuisance parameter is present only under the alternative", *Econometrica*, **62**, 1383-1414.

[2000] Ashley, R.A. and Patterson D.M. (2000) "Nonlinear model specification/diagnostics: Insights from a battery of nonlinearity tests" Department of Economics Working Paper E99-05: Virginia Polytechnic University.

[1997] Barnett W.A., Gallant A.R., Hinich M.J., Jungeilges J.A., Kaplan D.T. and Jensen M.J. (1997) "A single-blind controlled competition among tests for nonlinearity and chaos" *Journal of Econometrics*, **82**, 157-192.

[1988] Bose A. (1988) "Edgeworth correction by bootstrap an autoregressions" *The Annals of Statistics*, **16**, 1709-1722.

[1975] Brown R.L., Durbin J., Evans J.M. (1975) "Techniques for testing the constancy of regression relationships over time" (with discussion) *Journal of the Royal Statistical Society B*, **37**, 149-192.

[1999] Clements, M.P. and Hendry D.H. (1999) Forecasting Non-stationary Economic Time Series. MIT Press: Cambridge.

[1998] Dahl C.M. (1999) "An investigation of tests for linearity and the accuracy of flexible nonlinear inference", Working Paper 1999-8, Department of Economics, University of Aarhus.

[1987] Davies R.B. (1987) "Hypothesis testing when a nuisance parameter is only identified under the alternative" *Biometrika*, **47**, 33-43.

[1974] Engle, R. F. (1974) "Band Spectrum Regression" *International Economic Review,* **15**, 1-11.

[1996] Franses H. (1996) *Periodicity and Stochastic Trends in Economic Time Series*, Oxford University Press: Oxford.

[1991] Gallant A. R. and Souza, G.(1991) "On the asymptotic normality of fourier flexible form estimates" *Journal of Econometrics*, **50**, 329-53.

[1996] Hansen B. (1996) "Inference when a nuisance parameter is not identified under the null hypothesis" *Econometrica*, **64**, 413-430.

[1999] Hansen B. (1999) "Testing for linearity" *Journal of Economic Surveys*, **13**, 551-576.

[1989] Harvey A.C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press: Cambridge.

[1995] Hendry D.F. (1995) *Dynamic Econometrics*, Oxford University Press: Oxford.

[1995] Herwartz H. (1995) *Analyse saisonaler Zeitreihen mit Hilfe periodischer Zeitreihenmodelle*, Verlag Josef Eul: Bergisch-Gladbach.

[1993] Lee, T., White, H. and Granger, C.W.J., (1993) “Testing for neglected nonlinearity in time series models”, *Journal of Econometrics*, **56**, 269-290.

[1996] Li H. and Maddala G.S (1996) “Bootstrapping time series models” *Econometric Reviews*, **15**, 115-195 (incl. discussion).

[2000] Ludlow J. and Enders W. (2000) “Estimating non-linear ARMA models using fourier coefficients” *International Journal of Forecasting*, **16,** 333-347.

[1996] Lütkepohl H. and Herwartz H. (1996) “Specification of varying coefficient time series models via generalized flexible least squares”, *Journal of Econometrics*, **70**, 261-290.

[1989] Perron, P. (1989) “The Great Crash, the oil price shock and the unit root hypothesis” *Econometrica*, **57**, 1361-1401.

[1971] Rothenberg T.J. (1971) “Identification in parametric models” *Econometrica*, **39**, 577-591.

[1998] Stinchcombe M.B. and White H. (1998) “Consistent specification testing with nuisance parameters present only under the alternative” *Econometric Theory*, **14**, 295-325.

[1999] Tan H.B. and Ashley R. (1999) “Detection and modeling of regressioon parameter variation across frequencies with an application to testing the permanent income hypothesis” *Macroeconomic Dynamics,* 3, 69-83.

[1993] Teräsvirta T., Lin C.F. and Granger C.W.J. (1993) "Power of the neural network linearity test", *Journal of Time Series Analysis*, **14**, 209-220.