ELSEVIER

Journal of Econometrics 80 (1997) 325–353

JOURNAL OF Econometrics

Exact tests in single equation autoregressive distributed lag models

Jan F. Kiviet[a,*], Jean-Marie Dufour[b]

[a] *Tinbergen Institute and Faculty of Economics & Econometrics, University of Amsterdam, 1018 WB Amsterdam, Netherlands*

[b] *C.R.D.E. and Département de Sciences Economiques, Université de Montréal, Montréal Qué., Canada, H3C 3J7*

Abstract

For hypotheses on the coefficient values of the lagged-dependent variables in the ARX class of dynamic regression models, test procedures are developed which yield exact inference for given (up to an unknown scale factor) distribution of the innovation errors. They include exact tests on the maximum lag length, for structural change and on the presence of (seasonal or multiple) unit roots, i.e. they cover situations where usually asymptotic and non-exact *t*, *F*, AOC, ADF or HEGY tests are employed. The various procedures are demonstrated and compared in illustrative empirical models and the approach is critically discussed. © 1997 Elsevier Science S.A.

Key words: ARX model; Exact inference; Monte Carlo test; Order of dynamics; Structural change; Unit root
JEL classification: C22

1. Introduction

Maximum-likelihood-based inference procedures are generally very popular. This popularity is mainly based on asymptotic optimality properties and on computational convenience. Given present-day computer speed and facilities, however, practitioners can and should bring more aspects into their statistical utility function than just ease of computation and behavior in infinitely large samples. Nowadays a more challenging and appropriate objective is to employ procedures

* Corresponding author.

This work was supported by grants from the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and the Government of Québec (Fonds FCAR).

PII S0304-4076(97)00048-1

which optimize the actual efficiency and accuracy of inference from the finite set of sample data at hand.

As far as test procedures are concerned, it is no longer a serious requirement today for the userfriendliness of a test procedure that it has a null distribution (if only asymptotically) which can be tabulated. Instead, these days the profession can be much better served if provided with testing techniques accompanied with software for producing (exact) p-values of the relevant statistics. Whether or not the null distributions of the test statistics involved are invariant to given design characteristics is no longer a relevant issue. The prime objectives now for test procedures should be: (i) control over level in finite samples, the most central problem being the elimination of nuisance parameters; (ii) optimality properties under conceivable relevant and verifiable circumstances, or at least a good record in controlled experiments; and (iii) computational feasibility. Note also that a test procedure is meaningful only if the model under the null hypothesis is testable, i.e. if it is sufficiently restrictive to make the probability of certain non-trivial events (defined in terms of the available observations) boundable by a small probability (the level of the test) under all data-generating processes compatible with the null hypothesis. This is really a logical prerequisite of any testing exercise.

In Dufour and Kiviet (1993, 1996) these goals are pursued in the context of the simple first-order dynamic regression model by combining procedures put forward in Dufour (1989, 1990) and Kiviet and Phillips (1990, 1992). In the classic format of econometrics teaching this model usually generates the first encounter where the student is taught that (s)he may stick to the usual inference procedures (since they are still asymptotically valid under certain regularity conditions) although one has to accept that the magnitude of the approximation errors is unknown, and in principle unknowable, because it depends on the unknown values of the parameters of the model. Generally, bounds on such errors have not been established. Indeed, for certain problems (such as inference about long-run multipliers), the approximation error associated with any nuisance-parameter free asymptotic approximation must be arbitrarily bad on certain subsets of the parameter space; see Dufour (1996). Students are (still) taught just to live with this unpleasant situation and to find comfort in the thought that the larger the sample is, the smaller the committed errors will be. The relevant question how large a sample has to be in order to feel confident is usually left unanswered, or receives an answer that has no consequences for actual practice. That the committed errors may be quite substantial has been assessed in various Monte Carlo studies; see e.g. Kiviet (1985) and Nankervis and Savin (1987).

In this paper a considerably more general model is considered. We examine exact inference procedures for the one-equation higher-order autoregressive-distributed lag model, indicated by AD($p, p_1, \ldots, p_J$), represented by

$$y_t = \sum_{i=1}^{P} \lambda_i y_{t-i} + \sum_{j=1}^{J} \sum_{i=0}^{p_j} \delta_{ji} z_{t-i}^{(j)} + \varepsilon_t, \quad t = 1, \ldots, T, \tag{1}$$

where the regressors are finite-order distributed lags of the (lagged-) dependent variable and of J linearly independent regressors $z^{(j)}, j = 1,\ldots,J$. Using standard notation on polynomials in the lag- or backward-shift operator B, the model may also be expressed as

$$\lambda(B)y_t = \sum_{j=1}^{J}\delta_j(B)z_t^{(j)} + \varepsilon_t, \quad t = 1,\ldots,T, \tag{2}$$

where the order of the polynomial $\lambda(B)$ is p, and the order of the polynomials $\delta_j(B)$ is $p_j, j = 1,\ldots,J$. We examine inference procedures concerning the lagged-dependent regressor variable coefficients $\lambda_1,\ldots,\lambda_p$ only.

The goal of deriving operational finite-sample tests and confidence sets for the above model is achieved by combining three basic techniques. First, nuisance parameters (e.g., the unknown coefficients δ_{ji}) are eliminated by adding artificial regressors (at least when estimating the unrestricted model). The appropriate artificial regressors are obtained by extending the methods proposed in Kiviet and Phillips (1992) and Dufour and Kiviet (1993). This approach may also be interpreted as the one where we consider a class of invariant tests selected to eliminate the relevant nuisance parameters. Second, since the test statistics so obtained continue to have fairly complex null distributions which are difficult to compute analytically, even under a normality assumption on the errors (the Imhof algorithm cannot be applied here because one needs to compute distributions of multilinear forms in normal variables of order greater than two), we exploit the fact that the test statistics can be simulated easily and we use instead the technique of Monte Carlo tests originally suggested by Dwass (1957) and Barnard (1963). This technique replaces the original test proposed by a randomized analogue (so it is really a different test) which involves simulating the null distribution of the test statistic by Monte Carlo (MC) techniques. Although they are related to tests based on a parametric bootstrap, for which only large-sample justifications are available (see Efron and Tibshirani, 1993; Hall, 1992; Jeong and Maddala, 1993; and Vinod, 1993), MC tests have the important advantage of being provably valid in finite samples: irrespective of the number of replications used, which can be quite small (e.g., 19 artificial replications would be sufficient to obtain a test with a level of 0.05), the MC test has the right level. Furthermore, as the number of replications goes to infinity, it becomes equivalent to the original non-randomized test from which it was derived. For further general discussions of randomized tests, the reader may consult Lehmann (1986, Chapter 3), and for MC tests, Jöckel (1986) and Dufour (1995). Thirdly, we emphasize that finite-sample tests and confidence sets can be obtained relatively easily if we look first at the hypothesis that specifies the full vector of the autoregressive coefficients and then employ projection–intersection techniques (similar to those used in Dufour, 1989, 1990) to make inferences on individual coefficients or various parametric functions of interest (e.g., for testing a unit root hypothesis).

For further discussion of the importance of simultaneous inference (as opposed to marginal inference) in econometric models, see Dufour (1996).

The general outline of this study is as follows. In Section 2 we highlight the problem to be tackled and we state the assumptions which will make the above model genuinely testable. In Section 3 we derive an exact test for a joint hypothesis on all p elements of the vector $\lambda = (\lambda_1, \ldots, \lambda_p)'$. In Section 4 we examine tests on fewer than p restrictions on λ; we focus on tests for the order of the lag-polynomial $\lambda(B)$ and on tests for (multiple and/or seasonal) unit roots of $\lambda(B)$. In Section 5 we develop an exact test for the occurrence of structural change(s) in the values of the elements of the vector λ. Section 6 provides empirical illustrations of the various tests and makes comparisons with standard asymptotic results. Section 7 concludes and discusses the practical relevance of the results.

2. Framework

We may rewrite model (1) in matrix notation as

$$y = y\lambda + X\beta + \varepsilon, \tag{3}$$

where $Y = [Y_{.1} \vdots \cdots \vdots Y_{.p}]$ is a $T \times p$ matrix with $Y_{.i} = (y_{1-i}, \ldots, y_{T-i})'$ for $i = 1, \ldots, p$, and X is a $T \times k$ matrix with $k = J + \sum_{j=1}^{J} p_j$. The $k \times 1$ coefficient vector β contains all coefficients δ_{ji} in the appropriate order. From the particular temporal structure of the matrix Y it easily follows that, by introducing the $T \times T$ lower-triangula. matrix Γ and the $T \times p$ upper-triangular matrix Y_0, we may write

$$y - Y\lambda = \Gamma y - Y_0\lambda, \tag{4}$$

where

$$\Gamma = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ -\lambda_1 & 1 & 0 & & & & & & & \cdot \\ -\lambda_2 & -\lambda_1 & 1 & \cdot & & & & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & & & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & \cdot & & & & \cdot \\ -\lambda_p & & & \cdot & \cdot & \cdot & \cdot & & & \cdot \\ 0 & -\lambda_p & & & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & & \cdot & \cdot & 1 & 0 \\ 0 & \cdot & \cdot & 0 & -\lambda_p & \cdot & \cdot & -\lambda_2 & -\lambda_1 & 1 \end{bmatrix} \tag{5}$$

and

$$Y_0 = \begin{bmatrix} y_0 & y_{-1} & \cdot & \cdot & y_{1-p} \\ 0 & y_0 & \cdot & & y_{2-p} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & y_{-1} \\ \cdot & & & \cdot & y_0 \\ \cdot & & & & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} \tag{6}$$

Now we can rewrite model (3) as $\Gamma y = Y_0\lambda + X\beta + \varepsilon$, and since Γ is non-singular (irrespective of the actual value of λ), we have

$$y = \Gamma^{-1}Y_0\lambda + \Gamma^{-1}X\beta + \Gamma^{-1}\varepsilon \tag{7}$$

which, unlike (3), gives an explicit (reduced form) expression for all elements of the vector y. In our statistical analysis we shall treat both Y_0 and X as if they were fixed. More formally our framework is defined by:

Assumption A. The $T\times p$ and $T\times k$ regressor matrices Y and X in (3) *have* $\text{rank}([Y:X]) = p+k$ *with probability* 1. *The $T\times p$ matrix Y_0 given in* (6) *and the matrix X are both stochastically independent of the $T\times 1$ disturbance vector ε (strong exogeneity), and $\varepsilon = \sigma\eta$ with σ a scalar scale factor and η a $T\times 1$ vector with a known distribution. The parameters λ, β and σ are constant but unknown, with $\sigma \in \mathbb{R}^1_+, \beta \in \mathbb{R}^k, \lambda \in \mathscr{D}_\lambda \subseteq \mathbb{R}^p$, where $\mathscr{D}_\lambda$ is a priori specified.*

In other words, we suppose that the lagged-dependent variables in (3) capture the essential stochastic dynamic features of the model, as occuring, for example, in Dickey–Fuller regressions. To assume that feedbacks are possible between the error term ε_t and the regressors in X, even under a weak exogeneity assumption, would complicate considerably the relevant finite-sample theory (see, for example, Mankiw and Shapiro, 1986; Campbell and Dufour, 1991, 1995) and thus would require either modelling the joint behavior of all the variables involved (i.e., a multivariate model) or finding tests whose level is not affected by the form of the x_t process. Although certainly of interest, this goes well beyond the scope of the present paper.

Formula (7) enables us to represent the dependence of the lagged-dependent regressor variables $Y_{\cdot i}$ (and the full matrix of regressors Y) on fixed (observed and unobserved) and random (unobserved) elements. We have

$$Y_{\cdot i} = L^i y + Y_0 \iota_i, \quad i = 1, \ldots, p \tag{8}$$

where L is the $T\times T$ matrix

$$L=\begin{bmatrix}0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0\\ 1 & 0 & & & & & \cdot\\ 0 & 1 & 0 & & & & \cdot\\ \cdot & \cdot & \cdot & \cdot & & & \cdot\\ \cdot & & \cdot & \cdot & \cdot & & \cdot\\ \cdot & & & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & \cdot & \cdot & 0 & 1 & 0\end{bmatrix} \tag{9}$$

and ι_i denotes the $p\times 1$ unit vector with a unit element in the ith position, all others being zero. Substitution of (7) into (8) yields

$$\begin{aligned} Y_{\cdot i} &= L^i\Gamma^{-1}Y_0\lambda + Y_0\iota_i + L^i\Gamma^{-1}X\beta + \sigma L^i\Gamma^{-1}\eta\\ &= d_i(\lambda) + C_i(\lambda)X\beta + \sigma C_i(\lambda)\eta, \quad i=1,\ldots,p, \end{aligned} \tag{10}$$

where we introduced the notation

$$C_i(\lambda)=L^i\Gamma^{-1} \quad \text{and} \quad d_i(\lambda)=C_i(\lambda)Y_0\lambda + Y_0\iota_i. \tag{11}$$

From this decomposition of $Y_{\cdot i}$ in three terms the predetermined nature of $Y_{\cdot i}$ straightforwardly emerges: the first two components are fixed, and the third stochastic component is a linear transformation of the vector $\varepsilon=\sigma\eta$, where the transformation matrix $C_i(\lambda)$ is lower triangular with zeros on the main diagonal and also on $i-1$ lower subdiagonals. As will become clear below, it is the presence of the two fixed components in (10) that causes problems in finite samples when standard least-squares or (pseudo-) maximum-likelihood inference methods are applied to model (3).

We shall illustrate the problem with the standard procedure for the case of testing a joint hypothesis on all p coefficients in λ, i.e.

$$\mathrm{H}_0(\lambda):\ \lambda=\lambda_0 \text{ against } \mathrm{H}_1(\lambda):\ \lambda\neq\lambda_0, \tag{12}$$

where λ_0 is a p-element vector of known real numbers. In the standard test procedure ordinary least-squares estimation is performed and the F-statistic,

$$\mathcal{F}_\lambda=\frac{T-p-k}{p}\left[\frac{(y-Y\lambda_0)'M[X](y-Y\lambda_0)}{y'M[Y\vdots X]y}-1\right], \tag{13}$$

is employed (for any full column rank $T\times m$ matrix A, $M[A]=I_T-A(A'A)^{-1}A'$). Under $\mathrm{H}_0(\lambda)$ and usual regularity conditions

$$\mathcal{F}_\lambda \overset{a}{\sim} F(p,T-p-k). \tag{14}$$

Here $F(p, T-p-k)$ denotes the F-distribution with p and $T-p-k$ degrees of freedom respectively, and $\overset{a}{\sim}$ indicates the approximate (asymptotic) validity of statement (14). Even if η is multivariate standard normally distributed, the actual finite-sample distribution of statistic $\mathscr{F}_\lambda$ under $H_0(\lambda)$ is rather complicated and depends inter alia on the nuisance parameters β and σ, which in practice are unknown.

This non-similarity of the test can be seen as follows. We have

$$(y - Y\lambda_0)'M[X](y - Y\lambda_0)/\sigma^2 = \eta' M[X]\eta \quad \text{under } \mathrm{H}_0(\lambda). \tag{15}$$

When $\eta \sim N[0, I_T]$ this follows a $\chi^2(T-k)$ distribution. Complications arise, however, if in the minimization of the sum of squared residuals not all elements of λ are restricted as in the denominator of (13), where no constraints are imposed. Using well-known results from partitioned regression (see Dufour and Kiviet, 1993), we find

$$y'M[Y \vdots X]y = \sigma^2\eta' M[X]\eta - \sigma^2\eta' M[X]Y\{Y'M[X]Y\}^{-1}Y'M[X]\eta. \tag{16}$$

Hence, under $\mathrm{H}_0(\lambda)$,

$$\mathscr{F}_\lambda = \frac{T-p-k}{p}\left[\frac{\eta' M[X]\eta}{\eta' M[X]\eta - \eta' M[X]Y\{Y'M[X]Y\}^{-1}Y'M[X]\eta} - 1\right]. \tag{17}$$

From (10) we find that under $\mathrm{H}_0(\lambda)$

$$Y = \sum_{i=1}^{p} Y_{\cdot i}\iota_i' = \sum_{i=1}^{p}[d_i(\lambda_0) + C_i(\lambda_0)X\beta + \sigma C_i(\lambda_0)\eta]\iota_i' \tag{18}$$

from which it becomes obvious that $\eta' M[X]Y\{Y'M[X]Y\}^{-1}Y'M[X]\eta$, even under normality of η, has a distribution which is fairly complicated and depends on nuisance parameters. Y is stochastic, as it is determined by η, and (17) does not reduce to a simple ratio of quadratic forms in η, as in the standard case. Because of the presence of the two fixed components in (10) and (18) the scale factor σ does not drop out. Moreover, the distribution of (17) depends on the value of β. Both β and σ are unknown nuisance parameters when testing $\mathrm{H}_0(\lambda)$. Below, we consider an alternative to the test statistic $\mathscr{F}_\lambda$. Under $\lambda = \lambda_0$ this alternative will still have a complicated distribution function which is determined by λ_0 and X, but its distribution will be invariant to the unknown values β and σ, and hence enables one to produce exact inference on λ.

3. An exact joint test for the complete λ vector

The nuisance components revealed above can be removed by extending the model, and testing $\lambda = \lambda_0$ (or an extension of this null) in an appropriately

augmented model. We can obtain a pivotal test statistic by extending the set of strongly exogenous regressors X by some matrix $\tilde{X}$. This is chosen such that $[X \vdots \tilde{X}]$ has full column rank while $M[X \vdots \tilde{X}]$, unlike $M[X]$ in (17), annihilates all the nuisance components from Y. The p-values of such a statistic can be obtained by Monte Carlo methods. Such p-values are exact, even in a simulation analysis with a finite number of replications, when the test statistic is not considered in the usual fashion, but as a so-called Monte Carlo test, where the stochastic nature of both the test statistic and the sample of simulation drawings from its null distribution are taken into account. In the remainder of this section we show how to construct $\tilde{X}$ and how to obtain exact p-values.

From (18) and using (11), we find that the first nuisance component of Y is given by the $T \times p$ matrix

$$D_p(\lambda_0) = \sum_{i=1}^{p} d_i(\lambda_0) \iota_i' = \sum_{i=1}^{p} [C_i(\lambda_0) Y_0 \lambda_0 + Y_0 \iota_i] \iota_i'$$

$$= Y_0 + \sum_{i=1}^{p} C_i(\lambda_0) Y_0 \lambda_0 \iota_i' = Y_0 + \bar{C}_p(\lambda_0)[I_p \otimes Y_0 \lambda_0], \tag{19}$$

where

$$\bar{C}_p(\lambda_0) = [C_1(\lambda_0) \vdots \ldots \vdots C_p(\lambda_0)] \tag{20}$$

is a $T \times (pT)$ matrix. The dependence of the matrix $D_p(\lambda_0)$ on λ_0 is, of course, no problem since λ_0 is known. Extending the set of regressors $[Y \vdots X]$ of the model by including $D_p(\lambda)$, however, leads to removing the first nuisance component from the relevant test statistic.

It is less straightforward how to remove the second nuisance component of Y, because of its dependence on the unknown vector β. It equals

$$\sum_{i=1}^{p} C_i(\lambda_0) X \beta \iota_i' = [C_1(\lambda_0) X \beta \vdots \ldots \vdots C_p(\lambda_0) X \beta]$$

$$= \bar{C}_p(\lambda_0)[I_p \otimes X][I_p \otimes \beta]. \tag{21}$$

In order to annihilate this $T \times p$ component we have to project off the space spanned by the columns of the $T \times pk$ matrix

$$X_p(\lambda_0) = \bar{C}_p(\lambda_0)[I_p \otimes X] = [C_1(\lambda_0) X \vdots \ldots \vdots C_p(\lambda_0) X]. \tag{22}$$

Note that the $T \times (k + p + pk)$ matrix $[X \vdots D_p(\lambda_0) \vdots X_p(\lambda_0)]$ does not necessarily have full column rank. We therefore define the $T \times \tilde{m}$ matrix $\tilde{X}_p(\lambda_0)$ such that $[X \vdots \tilde{X}_p(\lambda_0)]$ has full column rank $m = k + \tilde{m}$, whilst the columns of the latter matrix span the same space as is spanned by the columns of the matrix

$[X \vdots D_p(\lambda_0) \vdots X_p(\lambda_0)]$. Consequently, if $\lambda = \lambda_0$, we have

$$M[X \vdots \tilde{X}_p(\lambda_0)]Y = \sigma \sum_{i=1}^{p} M[X \vdots \tilde{X}_p(\lambda_0)]C_i(\lambda_0)\eta \iota_i'$$

$$= \sigma M[X \vdots \tilde{X}_p(\lambda_0)]\bar{C}_p(\lambda_0)[I_p \otimes \eta] = \sigma Y(\eta, \lambda_0), \tag{23}$$

where the $T \times p$ matrix $Y(\eta, \lambda_0)$ is introduced to simplify the notation below.

Consider now testing $\lambda = \lambda_0$ in the extended model

$$y = Y\lambda + X\beta + \tilde{X}_p(\lambda_0)\tilde{\beta} + \sigma\eta, \tag{24}$$

where the $\tilde{m}$ regressors $\tilde{X}_p(\lambda_0)$ are redundant and hence actually $\tilde{\beta} \equiv 0$. We can proceed in two ways. We can either test

$$\tilde{H}_0(\lambda) : \lambda = \lambda_0 \quad \text{(leaving } \beta \text{ and } \tilde{\beta} \text{ unconstrained)} \tag{25}$$

or

$$\tilde{H}_0(\lambda, \tilde{\beta}) : \quad \lambda = \lambda_0, \quad \tilde{\beta} = 0. \tag{26}$$

For testing the p restrictions $\tilde{H}_0(\lambda)$ in model (24) we use statistic

$$\tilde{\mathscr{F}}_\lambda = \frac{T - p - m}{p} \left[\frac{(y - Y\lambda_0)'M[X \vdots \tilde{X}_p(\lambda_0)](y - Y\lambda_0)}{y'M[Y \vdots X \vdots \tilde{X}_p(\lambda_0)]y} - 1 \right]. \tag{27}$$

Under $\tilde{H}_0(\lambda)$ the first term in square brackets simplifies to

$$\frac{\eta' M[X \vdots \tilde{X}_p(\lambda_0)]\eta}{\eta' M[X \vdots \tilde{X}(\lambda_0)]\eta - \eta' Y(\eta, \lambda_0)\{Y(\eta, \lambda_0)'Y(\eta, \lambda_0)\}^{-1}Y(\eta, \lambda_0)'\eta}. \tag{28}$$

This is pivotal (it only depends on the fixed and known Y_0, X and λ_0, and on the stochastic vector η, which has a known distribution). For testing the $p + \tilde{m}$ restrictions $\tilde{H}_0(\lambda, \tilde{\beta})$ in (24) we use

$$\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}} = \frac{T - p - k - \tilde{m}}{p + \tilde{m}} \left[\frac{(y - Y\lambda_0)'M[X](y - Y\lambda_0)}{y'M[Y \vdots X \vdots \tilde{X}_p(\lambda_0)]y} - 1 \right]. \tag{29}$$

Under $\tilde{H}_0(\lambda, \tilde{\beta})$ the first term in square brackets is again pivotal. It is

$$\frac{\eta' M[X]\eta}{\eta' M[X \vdots \tilde{X}(\lambda_0)]\eta - \eta' Y(\eta, \lambda_0)\{Y(\eta, \lambda_0)'Y(\eta, \lambda_0)\}^{-1}Y(\eta, \lambda_0)'\eta}, \tag{30}$$

i.e. invariant with respect to β and σ, but dependent on X, Y_0 and λ_0.

Approximate critical- or p-values of tests (27) and (29) can be obtained from Monte Carlo experiments. However, exact p-values can easily be obtained from a finite number of simulation experiments for a closely related test procedure; see Dufour and Kiviet (1993, 1996). We illustrate this for (29).

First, generate $N-1$ (where $N=100$ or $N=1000$, for instance) independent T-element vectors η_j ($j = 1,\ldots,N-1$), which follow the same known distribution as $\eta = \varepsilon/\sigma$. From each vector η_j, a drawing $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(j)$ from the true null distribution of the test statistic can be obtained, i.e.

$$\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(j) = \frac{T-p-k-\tilde{m}}{p+\tilde{m}} \left[\frac{\eta_j' M[X]\eta_j}{\eta_j'\{M[X \vdots \tilde{X}(\lambda_0)] + M[Y(\eta_j,\lambda_0)] - I_T\}\eta_j} - 1 \right]. \tag{31}$$

To the series $\{\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(1),\ldots,\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(N-1)\}$, we add

$$\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(N) = \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}, \tag{32}$$

which represents the realization of the test statistic from the actual sample data. Next, we order the values $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(1),\ldots,\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(N)$ in increasing order. We indicate the position of $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(N)$ after this ordering by $\mathscr{R}(N)$. Of course, $\mathscr{R}(N)$ is a discrete random variable; the positive integers 1 through N are its domain, and since under the null all $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}(j)$ are i.i.d., each one of the possible outcomes has probability $1/N$. Instead of $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$, we use $\mathscr{R}(N)$ in the simulation or randomization test procedure as our test statistic. This is done in the following way. From our one and only realization of model (3), i.e. one realization of ε, we obtain one realization of statistic $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$, and, after generating a random series of $N-1$ Monte Carlo realizations of the null distribution (31), we have obtained one realization of the simulation test statistic, which we indicate by $\hat{\mathscr{R}}(N)$. Then the p-value of the simulation test statistic $\hat{\mathscr{R}}(N)$, which is based on $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$, is given by

$$\begin{aligned} \mathscr{P}(\lambda_0, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}) &= P[\mathscr{R}(N) \geq \hat{\mathscr{R}}(N) \mid \tilde{H}_0(\lambda,\tilde{\beta})] \\ &= \frac{N - \hat{\mathscr{R}}(N) + 1}{N}. \end{aligned} \tag{33}$$

Asymptotically ($N \to \infty$), this test procedure coincides with the procedure where statistic $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$ is used in combination with its exact null distribution. For finite N, we have a different procedure; its power will be related to the power of test $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$, but it is also characterized by N. However, even for finite N the randomized test is perfectly exact, since the probability to commit a type I error never exceeds the chosen significance level when we reject $\tilde{H}_0(\lambda,\tilde{\beta})$ if and only if $\mathscr{P}(\lambda_0, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}})$ is smaller than this level.

Although it is akin to a bootstrap (especially a parametric bootstrap), the above technique is conceptually different, as emphasized by Hall (1994, pp. 2342–2343). It was originally proposed by Dwass (1957) and Barnard (1963) as a method to perform *exact* tests based on statistics whose null distribution is intractable but can be simulated fairly easily. Bootstrap techniques were introduced much later (Efron, 1979) and, to date, have only an asymptotic ($T \to \infty$) justification (for reviews, see Efron and Tibshirani, 1993; Hall, 1992; Jeong and Maddala, 1993; Vinod, 1993). In a Monte Carlo test, the fundamental test one would like to perform is replaced by a randomized analogue, strictly speaking a different test, which has precisely the right level by a 'blurring' phenomenon (Marriott, 1979). This involves some power loss with respect to the (infeasible) fundamental test, although it is typically quite small (99 replications will usually be quite sufficient for a 5% level test), and the Monte Carlo test becomes equivalent to the fundamental test as the number of replications increases. For further discussion, see Dwass (1957), Marriott (1979), Jöckel (1986) and Dufour (1995).

Because of the validity of $\tilde{\beta} = 0$, the test $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$ seems logically the most attractive, and therefore may in general be more powerful than test $\tilde{\mathscr{F}}_{\lambda}$. Below, however, we will see that there is no uniform power difference between the two procedures. For the case $p = 1$ (first-order dynamics) the procedures developed above all simplify to variants of those presented in Dufour and Kiviet (1993), where, for two particular data sets, experimental evidence is provided that sustains the conjectured 'frequent superiority' of procedures that exploit $\tilde{\beta} \equiv 0$. In that study, it is also shown that exact procedures such as developed here may even be valuable in circumstances where the distribution of the disturbances is unknown or misspecified and where the regressors are weakly instead of strongly exogenous, since, under usual regularity assumptions, the suggested procedures will nonetheless be asymptotically valid and, although nonexact now, may often involve less serious finite-sample inaccuracies than the crude standard asymptotic inference techniques.

Finally, it is worthwhile noting here that the tests based on the statistics defined in (27) and (29) are both invariant to transformations $y_t \to y_{t*}, t = -p+1, \ldots, T$ such that

$$y_* - Y_*\lambda_0 = c(y - Y\lambda_0) + X\gamma, \tag{34}$$

where $c \neq 0$ and $\gamma \in \mathbb{R}^k$ are arbitrary constants [for another use of this invariance group in the context of a dynamic model, see Dufour and King (1991)]. More precisely, we consider a class of tests *invariant* with respect to the transformations (34) selected so that the null distribution of the test statistic does not depend on nuisance parameters. The latter feature is obtained by adding the artificial regressors $\tilde{X}_p(\lambda_0)$ to the regression (at least when estimating the model under the alternative). Since uniformly most powerful tests are not generally available in

autoregressive models (see Anderson, 1948), it is natural to consider a restricted class of tests such as a set of invariant tests. Because invariant test statistics have null distributions that do not depend on nuisance parameters, invariant tests may be interpreted as a subclass of *similar* tests. An alternative approach to eliminate nuisance parameters here would consist in characterizing directly the wider class of similar tests and then trying to find an optimal (or better) test in this class. This could be done, for example, by finding statistics that are *sufficient* for the nuisance parameters (at least under the null hypothesis) and then considering critical regions whose level is constant conditional on the sufficient statistics (tests with *Neyman structure*); see Lehmann (1986, Chapter 5) and Hillier (1987) for general descriptions and illustrations of this technique. In some cases, the required conditional distribution can be worked out relatively easily, but unfortunately this does not appear to be the case here even if one uses Monte Carlo techniques (we need to condition on events of zero probability). By contrast, as already emphasized in Durbin and Watson (1971) and Dufour and King (1991) for analogous test problems in dynamic models, invariant tests can be defined in terms of relatively simple transformations of the data (e.g., *maximal invariants*) and are thus quite easy to implement, especially when they are combined with the technique of MC tests.

4. Exact tests on special characteristics of the lag-polynomial $\lambda(B)$

In models with higher-order dynamics ($p>1$), we seldom want to test a hypothesis such as (12), where the complete vector λ is specified under the null. The most relevant example of this case is probably represented by $\lambda_0 \equiv 0$ for testing the presence of any lagged-dependent variables at all. Usually, however, we want to test less specific general characteristics of the lag-polynomial $\lambda(B)$, expressed by fewer than p linear restrictions.

A particularly relevant case is a test for the actual order of the lag-polynomial, i.e.

$$\lambda_p = \lambda_{p-1} = \cdots = \lambda_{p-r+1} = 0. \tag{35}$$

Here r zero restrictions ($1 \leqslant r \leqslant p$) are tested in order to check whether the order of the polynomial can be reduced from p to $p-r$. Another relevant case is the test for the presence of characteristic roots of the lag-polynomial of a particular value. Here the case of one or more (seasonal) unit roots deserves special attention. The single restriction that represents at least one unit root is

$$\lambda_1 + \lambda_2 + \cdots + \lambda_p = 1. \tag{36}$$

The presence of a second unit root implies the validity of the hypothesis:

$$\lambda_1 + 2\lambda_2 + \cdots + p\lambda_p = 0. \tag{37}$$

Our framework also extends to the seasonal unit root case. The assumption that the lag-polynomial $\lambda(B)$ can be factorized as

$$\lambda(B) = (1 - B^4)\lambda^*(B) = (1 - B)(1 + B)(1 + B^2)\lambda^*(B) \tag{38}$$

implies four linear restrictions on the λ coefficients. These are given by (36) and

$$-\lambda_1 + \lambda_2 - \lambda_3 + \cdots + (-1)^p \lambda_p = 1, \tag{39a}$$

$$\lambda_1 - \lambda_3 + \lambda_5 - \cdots = 0, \tag{39b}$$

$$-\lambda_2 + \lambda_4 - \lambda_6 + \cdots = 1. \tag{39c}$$

The latter two correspond to the two complex unit roots of (38), and (39a) follows from the minus unity root. Testing the four restrictions jointly or some (combination) of them separately using asymptotic methods is discussed in Dickey et al. (1984) and Hylleberg et al. (1990). Our procedures allow to produce exact tests for these various hypotheses.

Exact test procedures for r linear restrictions on the p elements of the coefficient vector λ can be obtained from the exact tests on the complete vector λ. Here we derive a general procedure, and we return to the various particular cases of interest given above in the illustrations in Section 6. As a building block for dealing with the case $r < p$, we need an exact confidence set with confidence coefficient $1 - \alpha$ for the complete coefficient vector λ. This is given by

$$\mathscr{C}_\lambda(\alpha, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}) = \left\{ \lambda_0 \in \mathscr{D}_\lambda : \mathscr{P}(\lambda_0, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}) \geqslant \alpha \right\}, \tag{40}$$

where, for the sake of brevity, we again only deal with the case where the test statistic $\tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}$ is being used. The dependence of the set on the value of N should be understood in the following broad sense: the calculation of the p-values $\mathscr{P}(\lambda_0, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}})$ is for each and every value λ_0 performed on the basis of the same set of randomly generated T-element vectors $(\eta_1, \ldots, \eta_{N-1})$.

The actual construction of set (40) for an empirical example requires extensive computations in order to establish to a prescribed degree of precision the boundaries of the confidence region. However, as it turns out, the test procedure for $r(< p)$ restrictions on $\lambda(B)$ developed below does not require the explicit construction of the p-dimensional set (40). We define a general representation of the r linear restrictions to be tested. Let R be a known $r \times p$ matrix with $\text{rank}(R) = r$, where $1 \leqslant r \leqslant p - 1$, and let

$$\theta = R\lambda. \tag{41}$$

Hence, θ is an $r \times 1$ vector of linearly independent linear transformations of λ. We want to devise an exact test procedure for

$$\text{H}_0(\theta): \theta = \theta_0 \quad \text{against} \quad \text{H}_1(\theta): \theta \neq \theta_0, \tag{42}$$

where θ_0 is a known r-element vector (for the case $r = 1$, we may also consider one-sided alternatives). For that purpose, we define the set

$$\mathscr{C}_\theta(\alpha, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}) = \left\{\theta_0 \in \mathbb{R}^r : \exists \lambda_0 \in \mathscr{C}_\lambda(\alpha, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}) \text{ such that } R\lambda_0 = \theta_0\right\}. \quad (43)$$

Obviously, we have

$$1 - \alpha = P[\lambda \in \mathscr{C}_\lambda(\alpha, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}})] \leqslant P[\theta \in \mathscr{C}_\theta(\alpha, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}})], \quad (44)$$

so that (43) establishes a conservative confidence set for θ. In order to test $H_0(\theta)$ of (42) for a specific known θ_0 value at level α one only has to check whether some $\lambda_0 \in \mathbb{R}^p$ exists which obeys the following two requirements: (i) $\lambda_0 \in \mathscr{C}_\lambda(\alpha, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}})$ and (ii) $R\lambda_0 = \theta_0$. $\mathrm{H}_0(\theta)$ is rejected when such a λ_0 value cannot be found; as soon as the search process establishes one λ_0 value that obeys the two requirements 'acceptance' of $\mathrm{H}_0(\theta)$ is legitimate. If $\mathrm{H}_0(\theta)$ and Assumption A are both true then the probability of rejection does not exceed the level α. To assess how conservative such a test is, it would be of interest to determine its size (as opposed to the level), i.e. the supremum of the rejection probability over all parameter vectors compatible with the null hypothesis, but this appears to be extremely difficult to do.

In practice, the search for a λ_0 value that satisfies the two requirements stated above is a computer-intensive problem. The procedure is relatively simple if r is large. For $r = p - 1$, the search is only over a straight line. If $r \ll p - 1$ [as it will often be, see (36)–(39)] more substantial numerical problems arise. Our approach employed in the empirical examples of Section 6 is as follows. Given the matrix R, we can find a $(p-r)\times p$ matrix $\bar{R}$ such that $R^+ = [R' \vdots \bar{R}']'$ is non-singular. Then we consider the reparametrization of λ to θ^+, where θ^+ is a $p\times 1$ vector:

$$\theta^+ = \begin{pmatrix} \theta \\ \bar{\theta} \end{pmatrix} = \begin{bmatrix} R \\ \bar{R} \end{bmatrix} \lambda = R^+ \lambda. \quad (45)$$

This reparametrization facilitates the search process since it enables one to impose the above-mentioned requirement (ii) directly, and makes the search solely over the space $\mathscr{D}_{\bar{\theta}}$ of the parameter $\bar{\theta} \in \mathscr{D}_{\bar{\theta}} \subseteq \mathbb{R}^{p-r}$, where $\mathscr{D}_{\bar{\theta}}$ is actually determined by $\mathscr{D}_\lambda$ and R, and may even be of smaller dimension than $p - r$.

Let the $(p-r)\times 1$ vector $\bar{\theta}_0^{(i)}$ be the ith value of $\bar{\theta}$ that is to be checked. So, the (in)validity of $\mathscr{P}(\lambda_0^{(i)}, N, \tilde{\mathscr{F}}_{\lambda,\tilde{\beta}}) \geqslant \alpha$ is verified, where

$$\lambda_0^{(i)} = (R^+)^{-1} \begin{pmatrix} \theta_0 \\ \bar{\theta}_0^{(i)} \end{pmatrix}. \quad (46)$$

Now the only problem left for executing the test is to devise a strategy for starting and updating the series $\{\bar{\theta}_0^{(i)}; i = 1, \ldots\}$, in combination with the design of a stopping criterion. This criterion is obvious for the case where an acceptable

$\lambda_0^{(i)}$ value has been found [this implies 'acceptance' of $H_0(\theta)$], but less so when all values of $\lambda_0^{(i)}$ according to (46) checked so far gave p-values smaller than α, and it looks rather unlikely that an acceptable value of $\lambda_0^{(i)}$, obeying the two requirements, will be found. In the implementation used in the empirical section below, we enhance the probability of finding an acceptable test value (if there are such in the $\mathscr{D}_{\bar{\theta}}$ space) straightaway, by taking for $\bar{\theta}_0^{(1)}$ the estimated value of $\bar{\theta}$ obtained when estimating

$$y = Y(R^+)^{-1}(R^+\lambda) + X\beta + \varepsilon = Y(R^+)^{-1}\theta^+ + X\beta + \varepsilon \tag{47}$$

under the r constraints $\theta = \theta_0 = R\lambda_0$.

5. Exact tests on structural change in the polynomial coefficients

Several finite-sample tests of parameter constancy against the presence of structural change for a linear regression model with one lagged-dependent variable are proposed in Dufour and Kiviet (1996). In that study we considered two distinct cases, depending on whether λ is assumed to be constant under the alternative or not. In the first case (λ fixed), we studied two categories of tests: analysis-of-covariance (AOC) type tests against alternatives where β may change at g known breakpoints, and CUSUM-type tests which are built against less specific alternatives. In the second case, we studied again two types of tests: predictive tests and AOC-type tests against alternatives where λ may change at one known breakpoint. For all these four types of tests we found exact procedures. They work satisfactorily, except for the last situation mentioned, i.e. the AOC change in λ test. All four structural change test procedures can now be generalized for the higher-order dynamic model by employing the results obtained in Section 3 of the present study. Since the focus of this paper is on inference regarding λ, and because the procedures developed for the higher-order model here suggest a simple alternative to the unsatisfactory (because of low power) change in λ test developed earlier (whilst the generalization to the $p > 1$ case is rather trivial for the other structural change tests), we shall only consider another and more promising AOC-type test on structural change in the p-element vector λ.

In Dufour and Kiviet (1996) we obtained a test of the constancy of λ for the case of one known breakpoint under the alternative by considering the difference between two particular (viz. invariant with respect to β and σ) estimates of λ, obtained from the two separate subsamples, respectively. Thus, it was avoided to obtain the test statistic from estimating the model from the full sample under the alternative hypothesis, which involves two regressors both including particular lagged-dependent variable observations. In the present study we have managed to deal with comparable multiple dynamic models, and hence it should now be possible to develop an AOC-type test.

We consider the case where there may by g ($\geqslant 1$) changes in the value of the parameter vector λ at g given breakpoints, viz. at

$$1 < T_1 < T_2 < \cdots < T_g < T. \tag{48}$$

We define

$$h(i) = \begin{cases} T_i & \text{for } i = 1,\ldots,g, \\ T+1 & \text{for } i = g+1, \end{cases} \tag{49}$$

and the $T\times T$ matrices H_t such that $H_1 = I_T, H_{T+1} = \mathrm{O}$ and

$$H_t = \begin{bmatrix} \mathrm{O} & \vdots & \mathrm{O} \\ \mathrm{O} & \vdots & I_{T-t+1} \end{bmatrix} \quad \text{for } t = 2,\ldots,T. \tag{50}$$

We also define the $T\times T$ matrices

$$G_i = H_{h(i)} - H_{h(i+1)}, \qquad \text{for } i = 1,\ldots,g, \tag{51}$$

the $T\times T(g+1)$ matrix

$$G = [I_T \vdots G_1 \vdots \ldots \vdots G_g], \tag{52}$$

and the $p(g+1)\times 1$ coefficient vector

$$\bar{\lambda} = \begin{bmatrix} \lambda \\ \lambda^{(1)} \\ \vdots \\ \lambda^{(g)} \end{bmatrix}, \tag{53}$$

where $\lambda^{(i)}, i = 1,\ldots,g$, are $p\times 1$ vectors. Now we generalize model (3) to

$$y = G(I_{g+1} \otimes Y)\bar{\lambda} + X\beta + \varepsilon. \tag{54}$$

This parametrization entails that over the period from observation 1 through $T_1 - 1$ the coefficient vector relating to the regressors Y is λ, from T_g through T this coefficient vector is $\lambda + \lambda^{(g)}$, whilst it is $\lambda + \lambda^{(i)}$ from T_i through $T_{i+1} - 1$ (where $i = 1,\ldots,g-1$).

For model (54) we shall consider tests of the hypothesis

$$H_0(\lambda^{(1)},\ldots,\lambda^{(g)}) : \lambda^{(1)} = \cdots = \lambda^{(g)} = 0, \tag{55}$$

which involves pg restrictions, and of

$$H_0(\bar{\lambda}) : \bar{\lambda} = \bar{\lambda}_0, \quad \text{where } \bar{\lambda}_0 = (\lambda_0', 0')', \tag{56}$$

which involves $p(g+1)$ restrictions, i.e. $\lambda = \lambda_0$ and $\lambda^{(1)} = \cdots = \lambda^{(g)} = 0$. This set-up concerns structural changes in the lagged-dependent variable coefficients only, but it is straightforward to allow for changes in (elements of) the β vector as well (under both the null and the alternative) by appropriately redefining X and β. For the sake of simplicity, we assume that the values T_i and the matrix $[Y \vdots X]$ are such that all elements of the coefficient vector $\bar{\lambda}$ are estimable, i.e. $[G(I_{g+1} \otimes Y) \vdots X]$ has full column rank.

$H_0(\bar{\lambda})$, given in (56), is a hypothesis which restricts all (transformed) lagged-dependent variable coefficients of (54), so in a properly augmented model exact tests can be constructed for the hypotheses $\tilde{H}_0(\bar{\lambda})$ and $\tilde{H}_0(\bar{\lambda}, \tilde{\beta})$. Analogous to what we did in Section 3, $\tilde{H}_0(\bar{\lambda})$ is tested by the statistic

$$\tilde{\mathscr{F}}_{\bar{\lambda}} = \frac{T - p(g+1) - \bar{m}}{p(g+1)} \left[\frac{(y - Y\lambda_0)' M[X \vdots \tilde{X}_p(\bar{\lambda}_0)](y - Y\lambda_0)}{y' M[G(I_{g+1} \otimes Y) \vdots X \vdots \tilde{X}_p(\bar{\lambda}_0)]y} - 1 \right], \quad (57)$$

which does not constrain the redundant regressor coefficients under the null, while for testing $\tilde{H}_0(\bar{\lambda}, \tilde{\beta})$ we use

$$\tilde{\mathscr{F}}_{\bar{\lambda},\tilde{\beta}} = \frac{T - p(g+1) - \bar{m}}{p(g+1) + \bar{m} - k} \left[\frac{(y - Y\lambda_0)' M[X](y - Y\lambda_0)}{y' M[G(I_{g+1} \otimes Y) \vdots X \vdots \tilde{X}_p(\bar{\lambda}_0)]y} - 1 \right], \quad (58)$$

which involves $\bar{m} - k$ extra zero restrictions on the redundant regressors. Here $\tilde{X}_p(\bar{\lambda}_0)$ is a $T \times (\bar{m} - k)$ matrix which is built as follows. Under $\tilde{H}_0(\bar{\lambda})$ we have

$$Y = D_p(\lambda_0) + X_p(\lambda_0)[I_p \otimes \beta] + \sigma \bar{C}_p(\lambda_0)[I_p \otimes \eta], \quad (59)$$

where the matrices $D_p(\lambda_0)$, $X_p(\lambda_0)$ and $\bar{C}_p(\lambda_0)$ are as defined in (19), (22) and (21), respectively. Hence, in order to obtain similarity of both test statistics, $[X \vdots \tilde{X}_p(\bar{\lambda}_0)]$ must have full column rank $\bar{m}$ and span the same column space as $[X \vdots G(I_{g+1} \otimes D_p(\lambda_0)) \vdots G(I_{g+1} \otimes X_p(\lambda_0))]$. Writing

$$\begin{aligned} Y(\eta, G, \lambda_0) &= \tfrac{1}{\sigma} M[X \vdots \tilde{X}_p(\bar{\lambda}_0)] G(I_{g+1} \otimes Y) \\ &= M[X \vdots \tilde{X}_p(\bar{\lambda}_0)] G \left[I_{g+1} \otimes [\bar{C}_p(\lambda_0))(I_p \otimes \eta)] \right], \end{aligned} \quad (60)$$

the denominator of the first term in square brackets in (57) and (58) can be written (after division by σ^2) as

$$\eta' M[X \vdots \tilde{X}_p(\bar{\lambda}_0)]\eta - \eta' Y(\eta, G, \lambda_0) \{ Y(\eta, G, \lambda_0)' Y(\eta, G, \lambda_0) \}^{-1} Y(\eta, G, \lambda_0)' \eta,$$

which shows that under $H_0(\bar{\lambda})$ the two test statistics are pivotal indeed, and can be used in a simulation test procedure to yield exact inference.

Testing $\tilde{H}_0(\bar{\lambda})$ or $\tilde{H}_0(\bar{\lambda}, \tilde{\beta})$ can be interesting, for instance, for the case $p=1$, $g=1$ and $\lambda_0 = 1$ (to check whether a unit root is valid also over the second part of a sample), but more generally – and similar to the procedures developed in the foregoing section – these tests are building blocks for exact (conservative) inference on (55), which does not constrain λ. This hypothesis is 'acceptable' if for some value λ_0 the hypothesis $\tilde{H}_0(\bar{\lambda})$ or $\tilde{H}_0(\bar{\lambda}, \tilde{\beta})$ is not rejected. If $\tilde{H}_0(\bar{\lambda})$ or $\tilde{H}_0(\bar{\lambda}, \tilde{\beta})$ is rejected for all λ_0 values located in a confidence set for $\bar{\lambda}$ then $H_0(\lambda^{(1)}, \ldots, \lambda^{(g)})$ should be rejected too.

6. Illustrations

For illustrative purposes we use a data set analyzed and published in Davidson and MacKinnon (1985), henceforth DM, containing Canadian quarterly data on housing starts (HS), gross national expenditure in 1971$ (Y) and real interest rates (RR). DM focus on testing the most appropriate functional form of the relationship between these variables, allowing for a simple dynamic specification. Here we shall adopt a particular functional form of the long-run equilibrium relationship [we use hs=ln(HS) as the dependent variable, and RR and y=ln(Y) as the explanatory variables], but focus on the most appropriate specification of the short-run dynamics. We shall use and compare (as far as possible) both the results of the usual asymptotic and the here developed finite-sample techniques.

Regression 1 is the model specification as preferred by DM (they remark, however, that it does not provide grounds for complacency!). Note that all three right-hand side variables are lagged one period. DM give an obvious economic explanation for that, and also remark that this makes the model much more useful for forecasting purposes. We add to this that if an adequate specification does not require contemporaneous explanatory variables indeed, this enhances the strength of the exogeneity characteristics of the regressor variables other than hs_{-1}, and, although it does not imply the strong exogeneity of RR_{-1} and y_{-1}, it will help to uphold (or approximate more closely) the validity of our Assumption A. On the other hand, however, we realize that the seasonal filter that has been applied (the available series were seasonally adjusted by a regression technique that allowed for time-varying seasonality) may undermine strong and even weak exogeneity. Since we use a smaller sample than DM, our results are slightly different (we withhold more initial observations in order to examine in a next stage higher-order dynamics). All standard asymptotic empirical results presented here have been obtained by version 7 of PcGive (see Doornik and Hendry, 1992). Estimates of the asymptotic standard errors of coefficient estimates are given in parentheses, and p-values based on asymptotic null distributions are given in

brackets. When these are smaller than the nominal level of 1% this is marked with two asterisks, and with one when smaller than 5%

Regression 1. Simple forecasting model for hs: 1956(1) to 1982(4)

$$\begin{aligned} \text{hs} = {} & 2.4768 + 0.6097\text{hs}_{-1} - 0.0411\text{RR}_{-1} + 0.1827 Y_{-1} \\ & (0.5380)\ (0.0700)\quad (0.0087)\quad (0.0547) \end{aligned}$$

$R^2 = 0.795$, $\hat{\sigma} = 0.1420$, RSS $= 2.0987$, $T = 108$

AR(1, 103) = 0.0940 [75.98%]	AR(3, 101) = 2.1487 [9.88%]
AR(2, 102) = 1.4453 [24.05%]	AR(4, 100) = 2.4660 [4.97%] *
ARCH(4, 96) = 0.8228 [51.38%]	$H(9,94) = 0.6228$ [77.49%]
$R(1, 103) = 0.0158$ [90.03%]	
$N(2) = 3.0495$ [21.77%]	⟨SK = −0.32; EK = 0.55⟩

Although the fit is not marvelous, all four coefficients appear highly significant. Most diagnostics approve this specification. Only tests for serial correlation (of fourth order or higher) in the disturbances expose weaknesses in this specification (possibly of the dynamics); AR(v_1, v_2) is an LM test against serial correlation of order v_1 , transformed such that its null distribution is approximately $F(v_1, v_2)$. ARCH(v_1, v_2) denotes a statistic that under the null hypothesis of no autoregressive conditional heteroscedasticity of order v_1 is approximately distributed as $F(v_1, v_2)$. $H(v_1, v_2)$ tests for correlation between the squared residuals and the regressors and their squares (and eventually cross-products), and under unconditional homoscedasticity it is approximately $F(v_1, v_2)$ distributed. $R(1, v_2)$ is the RESET test for significance of the squared-fitted values which should be compared with the $F(1, v_2)$ distribution, and $N(2)$ is asymptotically $\chi^2(2)$ distributed when the disturbances are normal; SK is the skewness, and EK the excess curtosis.

The dynamic misspecification of Regression 1 is apparent from estimation of a higher-order dynamic ARX-model, viz. AD(4,4,4), see Regression 2. Although the fit is still not spectacular, all diagnostics, except $N(2)$, are found to have high p-values, and the restrictions imposed by Regression 1 are strongly rejected (by the asymptotic methodology). It is noteworthy that (in congruence with the DM findings) the two current explanatory variables RR and y are both insignificant (when tested by asymptotic tests under the maintained hypothesis of weak exogeneity of all regressors), and that – despite the serious multicollinearity among these 15 regressors – 11 coefficients have absolute t-ratio's above unity and six above two. Note that the last four hypotheses tested fit precisely into the format

of Section 4. At the end of this Section we shall examine what inference can be obtained on the order of the polynomial $\lambda(B)$ by our exact procedures.

Regression 2. AD(4,4,4) *model for hs*: 1956(1) to 1982(4)

hs =	2.7888	+0.6544hs$_{-1}$	−0.0660RR$_{-1}$	+1.0744y_{-1}
	(0.6743)	(0.1004)	(0.0252)	(1.0653)
	−0.1643hs$_{-2}$	+0.2328hs$_{-3}$	−0.2452hs$_{-4}$	
	(0.1142)	(0.1172)	(0.0978)	
	+0.0086RR	+0.0507RR$_{-2}$	−0.0645RR$_{-3}$	+0.0289RR$_{-4}$
	(0.0166)	(0.0270)	(0.0264)	(0.0186)
	+0.4941y	+0.8558y$_{-2}$	−2.3794y$_{-3}$	+0.2444y$_{-4}$
	(0.9020)	(1.0343)	(0.9963)	(0.8625)

$R^2 = 0.842$, $\hat{\sigma} = 0.1321$, RSS=1.6251, $T = 108$

AR(1,92) = 0.4303 [51.35%]	AR(4,89) = 0.2581 [90.40%]
$N(2)$ = 5.8766 [5.30%]	⟨SK = −0.52; EK = 0.65⟩
F(3,90) = 0.1411 [93.51%]	⟨addition of seasonal dummies⟩
F(3,90) = 0.4974 [68.50%]	⟨additi on of fifth-order lags⟩
F(11,93) = 2.4637 [0.94%] * *	⟨reduction to Regression 1⟩
F(3,93) = 0.2128 [88.73%]	⟨omission of the regressors RR, y, y_{-4}⟩
F(1,93) = 6.2873 [1.39%]*	⟨reduction of polynomial $\lambda(B)$ to order 3⟩
F(2,93) = 3.4992 [3.43%]*	⟨reduction of polynomial $\lambda(B)$ to order 2⟩
F(3,93) = 2.7415 [4.76%]*	⟨reduction of polynomial $\lambda(B)$ to order 1⟩
F(4,93) = 13.723 [0.00%] * *	⟨reduction of polynomial $\lambda(B)$ to order 0⟩

From a conventional augmented Dickey–Fuller analysis all three variables show a univariate behavior that does not differ significantly from non-stationary processes with one unit root. In this respect, we focus on variable hs only. From Regression 3 we see that a sixth-order autoregression (in the usual differenced form with one lagged regressor in levels) with an intercept and no seasonal dummies nor a linear trend (both have extremely low F-values) yields an acceptable statistical representation of the dynamic structure. The series of tests on further reduction of the order of the dynamics performed fit completely into the framework of Section 4 and will be 'exactified' below. They involve zero restrictions on the lagged Δhs regressors; the first $F(6,95)$ test also restricts the coefficient of hs$_{-1}$ to unity. The last two tests mentioned for Regression 3 involve a zero restriction on regressor hs$_{-1}$ (i.e. a unit root). Therefore, the standard p-value presented for the second $F(6,95)$ test (this also reduces the order p to one) is unfounded (and therefore preceded by '!'), since the approximation of the asymptotic null

distribution by the F distribution is improper here. For testing the significance of hs_{-1}, the Dickey–Fuller critical values are asymptotically valid.

Regression 3. Univariate AR(6) *model for hs*: 1957(3) to 1982(4)

$$\begin{aligned}\Delta hs = {} & 1.0164 - 0.0951 hs_{-1} - 0.0830\,\Delta hs_{-1} - 0.1781\,\Delta hs_{-2} \\ & (0.5697)\quad (0.0534)\quad (0.0998)\quad (0.0967) \\ & +0.1034\,\Delta hs_{-3} - 0.1773\,\Delta hs_{-4} - 0.1664\,\Delta hs_{-5} \\ & (0.1004)\quad (0.0957)\quad (0.0988)\end{aligned}$$

$R^2 = 0.191$, $\hat{\sigma} = 0.1436$, RSS $= 1.95794$, $T = 102$

AR(1,94)	= 0.0781 [78.05%]	AR(4,91) = 0.4983[73.70%]
$N(2)$	= 7.5412 [2.30%]*	⟨SK = −0.57; EK = 0.78⟩
F(2,93)	= 0.2842 [75.32%]	⟨addition of two more lags⟩
F(1,95)	= 2.8396 [9.53%]	⟨reduction of polynomial $\lambda(B)$ to order 5⟩
F(2,95)	= 2.8165 [6.48%]	⟨reduction of polynomial $\lambda(B)$ to order 4⟩
F(3,95)	= 3.1541 [2.84%]*	⟨reduction of polynomial $\lambda(B)$ to order 3⟩
F(4,95)	= 3.3309 [1.34%]*	⟨reduction of polynomial $\lambda(B)$ to order 2⟩
F(5,95)	= 2.7848 [2.16%]*	⟨reduction of polynomial $\lambda(B)$ to order 1⟩
F(6,95)	= 54.513 [0.00%] * *	⟨reduction of polynomial $\lambda(B)$ to order 0⟩
F(6,95)	= 3.7402 [!0.22%]	⟨reduction to random walk with drift⟩
ADF	= −1.780 [> 10%]	

The procedures developed in the foregoing sections, which are to be employed now, produce p-values on any type of linear hypothesis on λ, including unit root cases. These are valid asymptotically (see Dufour and Kiviet 1993), and are even exactly correct in finite samples when the distribution of the disturbances is specified properly. To obtain empirical results the exact procedures have been programmed in the GAUSS 386 System. We first focus on tests on the dynamic specification and the presence of unit roots in the univariate model for hs. The first two blocks of results in Table 1 concern tests of various null hypotheses which fully specify ($r = 6 = p$) the $\lambda(B)$ polynomial of the model analyzed in Regression 3. F refers to the standard test (13) of p restrictions on $\lambda(B)$; the F^* and F^{**} columns refer to the statistics (27) and (29), respectively. These test the same restrictions in an augmented model, where F^{**} also imposes the zero restrictions on the redundant regressors. The table gives p-values which, for the asymptotic tests have been obtained with respect to the F distribution. (NB: This is only valid when the null hypothesis specifies the coefficients λ_0 such that all roots of $\lambda(B)$ are located outside the unit circle.) For the exact tests the p-values are obtained by confronting the test statistics with a series of $N - 1 = 499$ simulated drawings from their true null distribution. Unless stated otherwise, we

Table 1
Exact inference in the univariate model for hs of Regression 3 ($k = 1$ (intercept); $p = 6$; $r = 6$; $N = 500$; $\eta \sim N[0, I]$)

Null hypothesis							p-values (%)				
Elements of vector λ_0'							Asymptotic			Exact	
	$\lambda_{0,1}$	$\lambda_{0,2}$	$\lambda_{0,3}$	$\lambda_{0,4}$	$\lambda_{0,5}$	$\lambda_{0,6}$	F	F^*	F^{**}	F^*	F^{**}
$T = 102$: 1957(3) to 1982(4)											
(a)	0.822	−0.095	0.282	−0.281	0.011	0.166	100.0	99.6	99.7	99.8	99.6
(b)	1.0	0.0	0.0	0.0	0.0	0.0	!0.2	!0.5	!3.6	3.6	15.0
(c)	0.9	0.0	0.0	0.0	0.0	0.0	3.1	3.6	11.3	5.8	16.4
(d)	0.8	0.0	0.0	0.0	0.0	0.0	2.4	5.8	12.7	6.6	16.0
(e)	0.858	0.0	0.0	0.0	0.0	0.0	3.9	6.7	15.0	8.0	19.2
(f)	0.797	0.071	0.0	0.0	0.0	0.0	4.8	8.0	16.5	10.0	20.4
(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2
(h)	0.865	−0.090	0.287	−0.281	0.024	0.194	!78.6	!97.6	!96.9	99.8	99.6
(i)	1.205	−0.122	0.413	−0.436	0.093	−0.154	!0.0	!0.0	!0.0	0.2	1.2
(j)	0.665	−0.169	0.000	1.000	−0.665	0.169	!0.0	!0.0	!0.0	0.2	0.2
(k)	0.595	0.349	−0.135	0.166	−0.460	0.485	!0.0	!0.2	!1.7	4.4	11.6
$T = 52$: 1970(1) to 1982(4)											
(a)	0.822	−0.050	0.267	−0.407	0.092	0.190	100.0	99.4	78.6	100.0	89.6
(b)	1.0	0.0	0.0	0.0	0.0	0.0	!7.4	!2.3	!1.5	12.4	6.2
(c)	0.9	0.0	0.0	0.0	0.0	0.0	20.9	23.4	9.6	34.0	16.2
(d)	0.8	0.0	0.0	0.0	0.0	0.0	27.7	33.7	15.4	39.6	17.8
(e)	0.810	0.0	0.0	0.0	0.0	0.0	27.9	34.0	15.1	40.0	18.0
(f)	0.810	0.000	0.0	0.0	0.0	0.0	27.9	34.0	15.1	40.0	18.0
(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2
(h)	0.835	−0.038	0.277	−0.401	0.106	0.221	!99.9	!81.1	!60.4	96.6	86.0
(i)	1.222	−0.123	0.419	−0.583	0.291	−0.226	!0.3	!0.0	!0.0	1.0	1.4
(j)	0.728	−0.265	0.000	1.000	−0.728	0.265	!0.0	!0.0	!0.0	0.2	0.2
(k)	1.000	1.000	−0.121	0.038	−0.426	0.558	!10.1	!2.9	!6.5	21.6	27.0
$T = 102$; $T_1 = 51$: break at 1970(1)											
(a′)	0.795	−0.107	0.293	−0.182	−0.060	0.100	94.2	93.9	77.5	96.2	85.4
(b′)	1.0	0.0	0.0	0.0	0.0	0.0	!3.5	!0.2	!1.5	4.0	15.2
(k′)	0.595	0.349	−0.135	0.166	−0.460	0.485	!1.0	!0.1	!1.7	5.2	19.0

assume normality of the disturbances in the exact procedures. In the first block of the table we present results, labelled (a)–(k), for sample size $T = 102$ (as in Regression 3) and in the second block we analyze the same null hypotheses (a)–(k) for a subsample of size $T = 52$. The last three results in the third block regard structural change tests.

Just for curiosity we test in result (a) the 6 restrictions where λ_0, as given in the table, equals the least-squares estimate of λ. As expected, all p-values are equal or close to unity, especially for the larger sample. Result (b) fully relates to the (one but last) test in Regression 3, where the reduction of the model to a simple random walk with drift is examined, i.e. $\lambda_{0,1} = 1$ and $\lambda_{0,j} = 0$ for $j \geqslant 2$. Like the standard F test, also the asymptotic F^* and F^{**} tests seem to reject this hypothesis, but note that due to the unit root the asymptotic null distributions differ from F (therefore we marked the improper p-values again by '!'). We observe that the correct actual finite-sample p-values of the F^* and F^{**} tests lead to conflicting inferences at the 5% level. Contrary to our intuition the p-values of F^* are found to be smaller than those of F^{**} in the first block of the table.

We first discuss the results for the full sample and neglect, for the moment, the F^{**} results. So, from (b) we see that the exact F^* test rejects the random walk with drift specification. However, according to results (c) and (d), a stationary non-zero mean AR(1) process with an autocorrelation coefficient of 0.9 or 0.8, although rejected by the asymptotic F, is acceptable on basis of the exact procedure. The results (c) and (d), where r = 6, imply that conservative tests for the hypotheses $\lambda_6 = \cdots = \lambda_{6-r+1} = 0$, for $5 \geqslant r \geqslant 1$, do not lead to rejection either. For r=1,2 this is in agreement with the corresponding asymptotic results given in Regression 3, but we find conflict between the exact and asymptotic procedures for $3 \leqslant r \leqslant 5$. On the whole, we will see that the rejection probability of the asymptotic tests seems too high (i.e. the p-values are too low). Results (e) and (f) are specially tailored for producing inference on the order of the polynomial $\lambda(B)$, hence on fewer than p restrictions (although the null hypotheses actually tested do involve p restrictions). Here we used a λ_0 value obtained from (46), where θ_0 is chosen such that the r restrictions are imposed, and the $p-r$ elements of $\bar{\theta}_0$ are chosen equal to their least-squares estimate under the r restrictions; see (47). We argued in Section 4 that in that way an extensive search may be avoided because rejection of the p restrictions will be less likely than for most other possible values of $\bar{\theta}_0$. In (e) this strategy is followed to test the AR(1) specification and (f) tests the AR(2) model. We find a somewhat more marked acceptance of the AR(1) and the AR(2) specifications indeed, whereas the asymptotic F again produces rejections. In (g) we find a unanimous overall rejection of the omission of all lagged hs regressors.

Next we perform ADF-type exact unit root tests. In (h) initially the one unit root restriction (36) has been imposed, and the remaining $p-1$ restrictions resulted from constrained estimation. We see that the AR(6) model with one unit root is not rejected, which is in agreement with the asymptotic ADF result given

in Regression 3. In (i) restriction (37) is imposed as well. Hence, here we test for two unit roots. Hypothesis (i) is rejected, and we are inclined to believe that a further search through the appropriate $\mathscr{D}_0$-space will not lead to acceptance of the $I(2)$ hypothesis. Results (j) and (k) relate to the so-called HEGY seasonal unit roots tests, see Hylleberg et al. (1990). In (j) the λ_0 vector obeys the four restrictions (36) and (39a)–(39c). We find a clear rejection of the occurrence of a factor $(1 - B^4)$ in $\lambda(B)$ (which is no real surprise, given that hs is seasonally adjusted). In (k) the two restrictions (36) and (39a) are tested jointly and rejected by the exact F^* test.

In the second block of Table 1 we perform the same type of analysis, omitting the first 50 of the 102 sample data. We now see that the F^{**} p-values are lower [except in (k)] than the F^* values, as they usually are according to our experience. As far as we can see, the atypical results in this respect for the present full data set are just a rather exceptional case, due to the (accidental) relatively small residuals obtained for the first observations in this sample. This aspect is relevant because of the following. In all tests performed in the first two blocks of Table 1 we found $m = 7$. Hence, of the 12 redundant regressors only 6 had to be taken into account (we used a singular-value decomposition with a tolerance of 10^{-8}). It can be made plausible that in this simple model, where X consists of one column of unit elements only, the appropriate matrix $\tilde{X}$ consists of columns which are determined by a (linearly or nonlinearly) trended variable (heavily determined by the values of the elements of λ_0) and by the columns of the matrix $[I_p \vdots O]'$. Hence, since the redundant regressors are closely related to the dummy variables that would annihilate the contribution of the initial observations, they happen to have a minor effect in the present model on the obtained residual sum of squares, and therefore the F^{**} test is less powerful in the full sample. In contrast, we find in the subsample that the F^{**} exact p-values are not only smaller than the F^* values, but that they are often also smaller than the asymptotic Fp-values. This boosts our expectations regarding the relative power of the exact procedures.

In order to try the exact structural change tests developed in Section 5 we tested the constancy of the λ vector over the two separate parts of the sample. Note that $1 \leqslant \bar{m} \leqslant 25$. In result (a′) we consider the unrestricted AR(6) model. The chosen λ_0 value is the λ estimate obtained under the alternative specification. We see that constancy of λ is acceptable ($\bar{m} = 11$). In (b′) the null model is a simple random walk with drift over the full sample, and the alternative model is an AR(6) with a break in all λ coefficients at T_1 . Here the two exact tests give conflicting results ($\bar{m} = 9$). In (k') the λ_0 value is equal to (k) for $T = 102$. Hence, roots $+1$ and -1 are imposed. These were accepted for the second subsample and gave conflicting results over the full sample. Now, imposing them and testing for constancy, we find p-values above 5% ($\bar{m} = 14$).

To check the sensitivity of our exact results to the normality assumption, we also performed the simulation procedures under two alternative distributional as-

Table 2
Exact p-values (%) in the univariate model for hs of Regression 3 ($k = 1$ (intercept); $p = 6$; $r = 6$; $N = 500$)

$T = 52$: 1970(1) to 1982(4)					$T = 102$: 1957(3) to 1982(4)			
SK= −1; EK= 1.5			Cauchy		SK= −1; EK= 1.5		Cauchy	
	F^*	F^{**}	F^*	F^{**}	F^*	F^{**}	F^*	F^{**}
(a)	99.8	88.2	98.4	46.4	100.0	99.8	89.2	67.0
(b)	10.4	5.4	7.4	14.2	5.6	14.0	5.8	13.0
(c)	31.8	14.2	17.2	17.6	5.8	15.8	6.0	13.2
(d)	34.8	16.6	18.2	17.8	6.8	16.0	5.2	12.8
(e)	35.6	16.8	18.4	17.8	7.8	18.2	5.8	13.4
(f)	35.6	16.8	18.4	17.8	9.2	20.0	6.8	13.6
(g)	0.2	0.2	0.4	6.4	0.2	0.2	0.2	2.2
(h)	96.8	84.4	85.6	44.2	99.8	100.0	99.8	77.2
(i)	0.8	2.2	6.6	12.2	0.6	2.4	4.4	9.0
(j)	0.2	0.4	5.6	8.2	0.2	0.2	2.4	3.4
(k)	19.8	26.2	16.4	19.8	4.4	12.6	10.0	13.2

sumptions. We examined the changes in the p-values when it is assumed that the η vector consists of independent elements obtained from either the Cauchy distribution (which has no finite moments) or from a normalized and sign-changed χ^2 distribution. In the latter case we took $\eta_t = -(v_t - 8)/4$, where $v_t \sim \chi^2$ (8). It can be shown that this yields disturbances with SK= −1 and EK = 1.5 , which is not too far from the values actually obtained in the regressions. The results are collected in Table 2. For the transformed χ^2 disturbances we find slightly lower p-values for the smaller sample size; for T=102 the results are very close to the normal case. The effects of Cauchy disturbances are more pronounced. Often the F^* and F^{**} results are affected in an opposite way.

Finally, we employed various of the exact procedures in the context of an econometric model on housing starts. In Table 3 we examine tests on the order of the $\lambda(B)$ polynomial and a few other hypotheses, in a restricted version of the model of Regression 2 (the regressors RR, y and y_{-4} have been omitted). Here the maximum number of redundant regressors that may be required is 36, whereas actually only 12 had to be used. In result (A) we test again the values of the least-squares estimates and find a plausible result. In (B) we find conflicting evidence on the hypothesis that the polynomial $\lambda(B)$ can be reduced to $(1 - B)$. From (C) we see that we cannot omit all lagged-dependent variables, but (D), (E) and (F) show that reduction of the order of the $\lambda(B)$ polynomial to 1 is acceptable (which is in conflict with the results of asymptotic tests in the model with no redundant regressors; see also Regression 2). Result (G) indicates that a (1-B) factor in $\lambda(B)$ while maintaining p=4 is acceptable. In (H) we test for structural change. Now the maximum number of redundant regressors is 72,

Table 3
Exact inference in a model for hs (Regression 2 with 3 constraints) ($k = 8$; $p = 4$; $r = 4$; $N = 500$; $\eta \sim$ N[0,I], $T = 108$: 1956(1) to 1982(4) ; $T_1 = 57$: break at 1970(1))

Null hypothesis					p-values (%)				
Elements of vector λ_0'					Asymptotic			Exact	
	$\lambda_{0,1}$	$\lambda_{0,2}$	$\lambda_{0,3}$	$\lambda_{0,4}$	F	F^*	F^{**}	F^*	F^{**}
(A)	0.6649	−0.1641	0.2255	−0.2556	100.0	98.0	87.3	99.2	89.4
(B)	1.0000	0.0	0.0	0.0	!0.0	!0.0	!0.2	0.8	5.8
(C)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2
(D)	0.6463	−0.1576	0.0630	0.0	13.3	18.2	46.7	29.0	51.8
(E)	0.6407	−0.1187	0.0	0.0	11.5	13.3	43.0	20.4	49.4
(F)	0.5711	0.0	0.0	0.0	6.1	10.8	40.4	18.6	49.4
(G)	0.8735	−0.1275	0.3642	−0.1102	!0.0	!0.1	!2.4	17.0	27.0
(H)	0.7103	−0.2782	0.3111	−0.3462	56.0	61.0	30.5	71.2	37.4

whereas we actually had to use only 20. The chosen λ_0 value is the λ estimate obtained when a break is allowed for; constancy is not rejected.

In the original DM specification of Regression 1, which seems dynamically misspecified, further restrictions on $\lambda(B)$ should not be tested. We only checked whether structural change tests would expose the inadequacy of this model. This is not the case. Taking T_1=57, the relevant exact tests have p-values slightly above 50% , whilst the p-value of the asymptotic test on the significance of $\lambda^{(1)}$ is as disappointing, viz. 45% . This completes our use of these data on building starts. The reader should realize that these computations were not undertaken to throw new light on the house building industry, but only serve to illustrate the performance of new techniques for the econometric model building industry.

7. Conclusion and discussion

Standard test procedures used in econometrics to find an adequate specification of the short-run dynamics and long-run relationships in linear autoregressive distributed lag models typically only have an asymptotic justification and are non-exact in finite samples. Actual sample sizes are finite, and usually this size is fixed in the short run. Also, asymptotic large sample arguments, which may be very useful for the (dis)qualification of (in)consistent techniques in particular circumstances, cannot make actual sample sizes larger. To content oneself with the empirical results of asymptotically valid techniques just like that, usually implies that one accepts risks to draw false inferences that cannot be kept under control. In science in general, and in econometric statistical methodology in particular, such situations should be suppressed where possible.

Here we develop alternative inference procedures for single (reduced form) dynamic linear regression models which are exact under Assumption A, and we demonstrate their feasibility in illustrative empirical models. We only considered procedures for exact inference on the coefficients of the lagged-dependent variables. These, however, will allow to obtain exact inference results on almost any type of linear or nonlinear restriction on any of the parameters of such models, as is demonstrated in Dufour and Kiviet (1993) for the first-order dynamic model. The procedures make use of redundant regressor variables which annihilate nuisance parameter dependence and yield similarity of the test statistics. This allows us to use them as 'randomized tests', which implies that exact p-values can be obtained from a limited number of Monte Carlo replications. Critics on these exact techniques that may be put forward are, among other things, that:

(i) the addition of redundant regressors, although accomplishing full control over the level of the tests, will lead to power loss;
(ii) the requirement to adequately specify the distribution function of the disturbances cannot be fulfilled in practice since we lack (economic) theory on how to specify disturbance terms;
(iii) the requirement that the regressors, apart from the lagged-dependent variables, have to be strongly exogenous will not be fulfilled in most models of practical interest;
(iv) the simulation tests involve too much computational efforts.

Our response is as follows:

(i) Indeed we find that the standard asymptotic tests may (often but not systematically) have smaller p-values than the exact tests, but that does not mean that they have more power. Power can only be discussed if we know the size, and that is the basic weakness of the asymptotic methods: their actual finite-sample size depends on unknown parameters and the chosen nominal significance level is as firm on the actual size as the level of a pitfall. The p-values of the standard tests can only be judged after size correction. However, using the same sort of tests after addition of redundant regressors, and assessment of their true finite-sample null distribution via simulation, works in fact as a size correction. It certainly seems likely that more redundant regressors lead to less power. Therefore, it is comforting that we found that the actual number of redundant regressors required is much less than the theoretical number, due to linear dependencies. Moreover, from the results in Dufour and Kiviet (1993) we know that when more restrictions are tested fewer redundant regressors are required. Hence, if we do not test just λ but also elements of β, even fewer redundant regressors are required.

(ii) If the actual distribution of the disturbances is very non-normal (under usual regularity assumptions) then the accuracy of asymptotic methods, although asymptotically still correct, will be very poor in finite samples. Our methods allow to perform a sensitivity analysis of the exact results under various distributional

assumptions. Moreover, they possess the same asymptotic validity properties as the standard procedures (although the latter are even non-exact when the true distribution of the disturbances is known).

(iii) In case of weakly exogenous regressors we can again invoke the same asymptotic arguments as are used for the standard procedures. Our procedures are indeed only fully exact when the regressors are strongly exogenous. Although we agree that this is one of the weaker elements in the methods proposed here, the standard procedures suffer from the same weakness. Further research on the seriousness of this aspect in empirically relevant cases should be done (also including size and power comparisons). If the inaccuracies in inference which find their origin in conditioning are really serious, then it could be worthwhile to abandon conditioning and to model not only the endogenous, but also the weakly exogenous variables. Next, the present approach could be followed again to develop test statistics which are invariant with respect to the nuisance parameters in these enlarged models, thus allowing control over size. At the other end of the spectrum are the higher-order univariate AR models, possibly with an intercept, seasonal dummies and trends, which are used to characterize stationary and non- stationary (periodic) stochastic processes. These fit completely into our Assumption A so that exact versions of the ADF, HEGY and I(2) tests could be developed. Their relative performance should be further examined in controlled experiments.

(iv) The exact test statistics are, like the standard tests, simple ratios of residual sums of squares. To obtain the exact p-value, a relatively short series of simulated independent realizations of the test statistic has to be generated. Nowadays this requires only a few lines of computer code and each application takes just a few seconds or less.

References

Anderson, T.W., 1948. On the theory of testing serial correlation. Skandinavisk Aktuarietidskrift 31, 66–116.

Barnard, G.A., 1963. Comment on 'The spectral analysis of point processes' by M.S. Bartlett. Journal of the Royal Statistical Society Series B 25, 294.

Campbell, B., Dufour, J.-M., 1991. Over-rejections in rational expectations models: A nonparametric approach to the Mankiw-Shapiro problem. Economics Letters 35, 285–290.

Campbell, B., Dufour, J.-M., 1995. Exact nonparametric orthogonality and random walk tests. Review of Economics and Statistics 77, 1–16.

Davidson, R., MacKinnon, J.G., 1985. Testing linear and loglinear regressions against Box-Cox alternatives. Canadian Journal of Economics 18, 499–517.

Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49, 1057–1072.

Dickey, D.A., Hasza, D.P., Fuller, W.A., 1984. Testing for unit roots in seasonal time series. Journal of the American Statistical Association 79, 355–367.

Doornik, J.A., Hendry, D.F., 1992. PcGive version 7: an interactive econometric modelling system. Institute of Economics and Statistics, University of Oxford.

Dufour, J.-M., 1989. Non-linear hypotheses, inequality restrictions, and non-nested hypotheses: exact simultaneous tests in linear regressions. Econometrica 57, 335–355.

Dufour, J.-M., 1990. Exact tests and confidence sets in linear regressions with autocorrelated errors. Econometrica 58, 479–494.

Dufour, J.-M., 1996. Some impossibility theorems in econometrics, with applications to structural and dynamic models. Econometrica (forthcoming)

Dufour, J.-M., 1995. Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics in econometrics. Technical Report (Centre de Recherche et Développement en Economique, Université de Montréal).

Dufour, J.-M., King, M.L., 1991. Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors. Journal of Econometrics 47, 115–143.

Dufour, J.-M., Kiviet, J.F., 1993. Exact inference methods for first-order autoregressive distributed lag models, Paper Presented at ESEM-93, Uppsala, Sweden (TI discussion paper, 94–143) forthcoming in Econometrica.

Dufour, J.-M., Kiviet, J.F., 1996. Exact tests for structural change in first-order dynamic models. Journal of Econometrics 70, 39–68.

Durbin, J., Watson, G.S., 1971. Testing for serial correlation in least squares regression III. Biometrika 58, 1–19.

Dwass, M., 1957. Modified randomization tests for nonparametric hypotheses. Annals of Mathematical Statistics 28, 181–187.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability, Chapman & Hall, Vol. 57, New York.

Hall, P., 1992. The Bootstrap and Edgeworth Expansion. Springer, New York.

Hall, P., 1994. Methodology and theory for the bootstrap. in: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, Volume IV Elsevier, Amsterdam, pp. 2341–2381.

Hillier, G.H., 1987. Classes of similar regions and their properties for econometric testing problems. Econometric Theory 3, 1–44.

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

Jeong, J., Maddala, G.S., 1993. A perspective on application of bootstrap methods in econometrics. In: Maddala, G.S., Rao, C.R., Vinod, H.D. (Eds.), Handbook of Statistics, Vol. 11, Econometrics North-Holland, Amsterdam, 573–610.

Jöckel, K.-H., 1986. Finite sample properties and asymptotic efficiency of Monte Carlo tests. Annals of Statistics 14, 336–347.

Kiviet, J.F., 1985. Model selection test procedures in a single linear equation of a dynamic simultaneous system and their defects in small samples. Journal of Econometrics 28, 327–362.

Kiviet, J.F., Phillips, G.D.A., 1990. Exact similar tests for the root of a first-order autoregressive regression model, University of Amsterdam. AE Report 12/90, Paper presented at ESWM'90, Barcelona.

Kiviet, J.F., Phillips, G.D.A., 1992. Exact similar tests for unit roots and cointegration. Oxford Bulletin of Economics and Statistics 54, 349–367.

Lehmann, E.L., 1986. Testing Statistical Hypotheses, 2nd ed. Wiley, New York.

Mankiw, N.G., Shapiro, M.D., 1986. Do we reject too often? Small sample properties of rational expectation models. Economics Letters 20, 243–247.

Marriott, F.H.C., 1979. Barnard's Monte Carlo tests: How many simulations? Applied Statistics 28, 75–77.

Nankervis, J.C., Savin, N.E., 1987. Finite sample distributions of t and F statistics in an AR(1) model with an exogenous variable. Econometric Theory 3, 387–408.

Vinod, H.D., 1993. Bootstrap methods: Applications in econometrics. In: Maddala, G.S., Rao, C.R., Vinod, H.D. (Eds.), Handbook of Statistics, Vol. 11, Econometrics, North-Holland, Amsterdam. pp. 629–661.