Econometrica, Vol. 66, No. 1 (January, 1998), 79–104

EXACT INFERENCE METHODS FOR FIRST-ORDER AUTOREGRESSIVE DISTRIBUTED LAG MODELS

BY JEAN-MARIE DUFOUR AND JAN F. KIVIET[1]

Methods are proposed to build exact tests and confidence sets in the linear first-order autoregressive distributed lag model with i.i.d. disturbances. For general linear hypotheses on the regression coefficients, inference procedures are obtained which have known level. The tests proposed are either similar (i.e., they have constant rejection probability for all data generating processes consistent with the null hypothesis) or use bounds which are free of nuisance parameters. Correspondingly the confidence sets are either similar with known size (i.e., they have constant coverage probability) or conservative. We also develop exact tests and confidence sets for various nonlinear transformations of model parameters, such as long-run multipliers and mean lags. The practical usefulness of these exact methods, which are also asymptotically valid under weak regularity conditions, is illustrated by some power comparisons and with applications to a dynamic trend model of money velocity and a model of money demand.

KEYWORDS: Autoregression, exact inference, general restrictions, Monte Carlo test, nuisance parameter elimination, similar test.

1. INTRODUCTION

A SERIOUS PROBLEM IN ECONOMETRICS is that statistical procedures which yield exact finite-sample inference in simple static linear models have an indeterminate distribution in more general and realistic models, due to the effects of unknown nuisance parameters. In the context of dynamic models, such problems do not occur when the relationship can be modeled according to the normal classical linear model with finite distributed lags. However, if the relationship has infinite distributed lags and is modeled in the popular autoregressive form, then the nuisance parameter problem arises. Since invariance of standard inference techniques with respect to nuisance parameters is often regained asymptotically as the sample size increases, it is common practice to employ asymptotic approximations. However, by simply relying on first-order asymptotic results when analyzing a finite (and often rather small) sample, one accepts to commit approximation errors of a largely unknown nature and magnitude. This objectionable practice is due to the fact that hardly any operational exact inference procedures are available to date. Hillier (1987) characterizes the problems involved and provides solutions to some particular testing problems. Here we develop operational exact inference techniques for (any transformation of) all coefficients of autoregressive models with exogenous regressors, whether stable (stationary) or not.

[1] The authors thank a co-editor and three anonymous referees for several useful comments. 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). An earlier version (see Dufour and Kiviet (1994)) has been presented at the Econometric Society European Meeting in Uppsala (Sweden), August, 1993.

In the literature some solutions to the nuisance parameter problem have been suggested for very simple dynamic models. Most attention has been paid to the model where the lags are characterized by the presence of just one common factor (see Sargan (1980)) and where all the explanatory variables are fixed. This linear regression model with AR(1) errors can be stated as:

$$(1.1) \qquad y_t = x_t'\beta + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{IIN}(0, \sigma^2) \qquad (t = 1, \ldots, T),$$

where y_t is the dependent variable, x_t is a $k \times 1$ vector of fixed (or strongly exogenous) regressors, β is a $k \times 1$ vector of fixed coefficients, and u_t is a random disturbance; the parameters β, ρ, and σ^2 are unknown. The central difficulty here comes from the fact that the distributions of test statistics (for hypotheses about β, for example) usually depend on unknown nuisance parameters (ρ, for example). To keep on using the same statistics requires one to bound their null distributions over the nuisance parameter space. For some illustrations of this approach, see Hillier and King (1987), Kiviet (1980, 1991), Vinod (1976), Vinod and Ullah (1981, Ch. 4), and Zinde-Walsh and Ullah (1987). The major drawback here is that the appropriate bounds on critical values may be very large, if not infinite; see Krämer, Kiviet, and Breitung (1991). An alternative bounding technique, which does not raise the same difficulties and involves using modified test statistics, has been suggested in Dufour (1990). This method starts from an exact confidence set for ρ, constructed by "inverting" an exact autocorrelation test, to obtain a *simultaneous* confidence set for ρ and any element of β. Corresponding exact similar tests follow. Then, using a projection method, exact confidence sets are constructed for the individual components of β from which bounds-type tests for hypotheses on these components are obtained. Because the confidence set on ρ actually restricts the nuisance parameter space and more efficient (ρ-dependent) test statistics are used, this method avoids the main problems associated with other bounds methods.

It should be recognized that model (1.1), which is one of the simplest dynamic alternatives to a fully static model (where $\rho = 0$ and x_t does not contain any lagged explanatory variables), has serious shortcomings when the relationship is genuinely dynamic; see Hendry and Mizon (1978). Here, we develop exact inference procedures for the coefficients of the model:

$$(1.2) \qquad y_t = \lambda y_{t-1} + x_t'\beta + \varepsilon_t, \quad \varepsilon_t \sim \text{IID}(0, \sigma^2) \qquad (t = 1, \ldots, T),$$

where y_t, x_t, ε_t, and β are defined as in (1.1). No assumptions are made on the domain of λ or the distribution of y_0. The exact inference procedures to be developed are based on test statistics whose distributions, under the null hypothesis, do not depend on nuisance parameters (λ, β, or σ) or the distribution of y_0; hence y_0 may either be fixed or random. Normality of the disturbances is not required for the invariance properties to hold. However, the actual distribution function has to be known for the calculation of exact significance points, and this is often relatively easy under normality. Note that model (1.2) includes as a special case the simple Dickey-Fuller type model for inference on one unit root in the presence of an intercept, drift, polynomial trends, seasonal dummies, or any interactions.

The basic building block that is exploited in Dufour (1990) to obtain exact inference for any subvector of β in model (1.1) is an exact confidence interval for the parameter ρ. In the same way an exact confidence interval for the coefficient λ would enable exact inference in the more general dynamic model (1.2). Exact inference on λ can be obtained by applying least-squares to an augmented regression model, as set out in Kiviet and Phillips (1990, 1992). Here we first develop a related but more general procedure for testing a joint hypothesis on λ and any linear transformation of β. This procedure has a neat likelihood-ratio (LR) interpretation and is therefore less ad hoc than earlier procedures. As a possibly simpler alternative, we also show that confidence sets can be built using the two-stage approach proposed in Dufour (1990), which combines an exact confidence interval for λ with the corresponding family of "conditional" confidence sets for the relevant linear transform of β. Then exact inference for any vector linear transformation (or subvector) of β is obtained by applying a union-intersection method.

Critical values or p values for the test statistics mentioned above can be obtained by Monte Carlo experiments on particular pivotal functions whose distributions depend on the parameters specified by the null hypothesis and the adopted distribution type for the disturbances. In principle, critical points or p values may be evaluated to any desired degree of precision by simulation. However, the number of required replications may get extremely large. Therefore, we propose to use instead the technique of Monte Carlo tests (Dwass (1957), Barnard (1963)) which allows one to obtain genuinely exact tests from any given (possibly very small) number of replications.

Apart from inference on linear transformations of the coefficients of model (1.2), practitioners are usually interested in inference on particular nonlinear transformations, such as long-run multipliers, (interim) impact multipliers, or mean and median lags. Exact procedures for testing nonlinear hypotheses in the static linear model are examined in Dufour (1989). We show here that analogous procedures can be applied in the dynamic model (1.2) in a way that exact inference on any nonlinear hypothesis can be obtained.

The structure of this paper is as follows. In Section 2, we develop procedures for exact joint inference on the value of λ (the coefficient of the lagged dependent variable) and linear transformations of β (the coefficients of the exogenous regressors) by putting the test problem into a form such that the LR-type test statistic does not depend on nuisance parameters. From these, exact inference procedures on λ only are also derived. The latter are then used in Section 3 for testing hypotheses which do not involve λ at all (i.e., which only restrict β) or do not involve λ explicitly. In Section 4, we develop exact tests on nonlinear transformations of the coefficients such as long-run multipliers. In Section 5, we make the various inference procedures operational by putting them into a Monte Carlo testing framework. Section 6 compares through Monte

Carlo experiments the exact tests with asymptotic (size corrected) procedures, while Section 7 illustrates the practical usefulness of the exact inference techniques by applying them to a dynamic trend model of money velocity in the U.S. and a model of Canadian money demand. Section 8 concludes.

2. EXACT JOINT INFERENCE ON λ AND β

Due to its notation, model (1.2) has the appearance of a simple partial adjustment model. However, the relationship may involve dynamics of a more sophisticated nature, such as

$$y_t = \lambda y_{t-1} + \sum_{j=1}^{J} \sum_{i=0}^{L(j)} \delta_{ji} z_{t-i}^{(j)} + \varepsilon_t \qquad (t = 1, \ldots, T). \tag{2.1}$$

In matrix form, the model we study can be written

$$\mathbf{y} = \lambda \mathbf{y}_{-1} + X\beta + \varepsilon, \tag{2.2}$$

where $X = [x_1, \ldots, x_T]'$ is a $T \times k$ matrix, and $\mathbf{y} = (y_1, \ldots, y_T)'$, $\mathbf{y}_{-1} = (y_0, \ldots, y_{T-1})'$, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ are $T \times 1$ vectors.

To obtain finite sample results, we make the following assumptions.

ASSUMPTION A: *The $T \times k$ matrix X and the $T \times 1$ vector ε are independent, y_0 is independent of ε, and rank$(X) = k$ with probability* 1; *$\beta \in \mathbb{R}^k$ and $\lambda \in \mathscr{D}_\lambda$ are fixed but unknown coefficients, with $\mathscr{D}_\lambda = \{\lambda \in \mathbb{R} \mid -\infty \leq \lambda_L \leq \lambda \leq \lambda_R \leq \infty\}$.*

ASSUMPTION B: *The distribution of the vector ε/σ (given X and y_0) is known, where σ is an unknown positive constant.*

Several of the invariance results given below hold without Assumption B. The assumption that X and ε are independent (*strong exogeneity*) entails that X can be treated as fixed for inference purposes, which we shall do from now on. Assumption B means that the distribution of ε is known up to a scale factor. Often we will suppose that the elements of ε are i.i.d. normally distributed with mean zero, but it will be straightforward to use other distributions, provided they are known up to a scale factor. This also includes cases where the elements of ε show heteroscedasticity and/or serial correlation of a given form, so that the model can be transformed to the form (2.2).

We first consider tests on the value of λ jointly with r linear restrictions on β. From these, procedures for inference on λ only will follow upon taking $r = 0$. Let R be a known $r \times k$ matrix with rank$(R) = r$. Since R may be void, we have $0 \leq r \leq k$. When $r \geq 1$, we consider

$$\gamma = R\beta. \tag{2.3}$$

To keep notation simple, we reparameterize the model and transform the regressors. Let $\bar{R}$ be a $(k - r) \times k$ matrix, such that $Q = [R' \,\vdots\, \bar{R}']'$ is nonsingular, and let $Q\beta = (\gamma', \bar{\gamma}')'$; of course, $\bar{R}$ and $\bar{\gamma} = \bar{R}\beta$ are not unique. Equation (2.2) may now be rewritten as

$$(2.4) \qquad \mathbf{y} = \lambda \mathbf{y}_{-1} + XQ^{-1}Q\beta + \varepsilon = \lambda \mathbf{y}_{-1} + Z\gamma + \bar{Z}\bar{\gamma} + \varepsilon,$$

where $[Z \,\vdots\, \bar{Z}] = XQ^{-1}$ has full column rank. When $r = 0$, we set $Q = \bar{R}$, $Q\beta = \bar{\gamma}$, $\bar{Z} = XQ^{-1}$, and $Z\gamma$ is simply dropped from (2.4).

Below, various tests for the joint hypothesis $\lambda = \lambda_0$, $\gamma = \gamma_0$ are examined. However, sometimes we will extend the null hypothesis and we may also test it in a model that differs from (2.4). Therefore, we introduce a notation for null hypotheses and test statistics which uniquely refers to both the restrictions tested and the model employed to test them:

(2.5) $\quad H(\lambda_0, \gamma_0)$ is the null hypothesis $H_0: \lambda = \lambda_0, \gamma = \gamma_0$ in model (2.4).

Conditional on y_0 and assuming the disturbances ε_t are i.i.d. normal, the LR statistic for testing $H(\lambda_0, \gamma_0)$ is $T \cdot \ln(S_0[\bar{Z} | \lambda_0, \gamma_0]/S_1[W])$, where

$$(2.6) \qquad S_0\left[\bar{Z} | \lambda_0, \gamma_0\right] = (\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)' M[\bar{Z}](\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0),$$

$$(2.7) \qquad S_1[W] = \mathbf{y}' M[W] \mathbf{y} = \varepsilon' M[W] \varepsilon,$$

are respectively the restricted and unrestricted sum of squared residuals, where $W = [y_{-1} \,\vdots\, X]$. We define $M[A] \equiv I - A(A'A)^{-}A'$ for any (possibly non-full column rank) matrix A. Note that $M[\mathbf{y}_{-1} \,\vdots\, X] \equiv M[\mathbf{y}_{-1} \,\vdots\, XQ^{-1}]$. We now consider the following monotonic transformation of the LR statistic:

$$(2.8) \qquad \mathscr{L}_{\lambda, \gamma}(\lambda_0, \gamma_0) = S_0\left[\bar{Z} | \lambda_0, \gamma_0\right] / S_1[W].$$

Under $H(\lambda_0, \gamma_0)$ we have $S_0[\bar{Z} | \gamma_0, \lambda_0] = \varepsilon' M[\bar{Z}] \varepsilon$. Further, from the partition $W = [y_{-1} \,\vdots\, X]$ and applying the Frisch-Waugh theorem (see Gouriéroux and Monfort (1989, Vol. 2, Rappel R.E.3, p. 478)) to $M[W]$, we can also write[2]

$$(2.9) \qquad S_1[W] = \varepsilon' M[W] \varepsilon = \varepsilon' M[X] \varepsilon - \left[(\varepsilon' M[X] \mathbf{y}_{-1})^2 / \mathbf{y}'_{-1} M[X] \mathbf{y}_{-1}\right].$$

From (2.9), we see easily that the statistic $\mathscr{L}_{\lambda, \gamma}(\lambda_0, \gamma_0)$ is not pivotal under the null hypothesis, i.e., its distribution depends on unknown nuisance parameters. This follows on observing that the numerator $\varepsilon' M[\bar{Z}] \varepsilon$ of $\mathscr{L}_{\lambda, \gamma}(\lambda_0, \gamma_0)$ and its denominator (2.9) can be divided by σ^2, and upon recognizing that the test statistic then consists of a number of (inner) products of three stochastic vectors. The vectors $M[\bar{Z}]\varepsilon/\sigma$ and $M[X]\varepsilon/\sigma$ are pivotal, but $M[X]\mathbf{y}_{-1}/\sigma$ is not, even under the null: from (2.2), we have

$$(2.10) \qquad \mathbf{y}_{-1} = y_0 \iota(\lambda) + C(\lambda) X\beta + C(\lambda)\varepsilon,$$

[2] More explicitly, for any full column rank matrix C, let $P[C] = C(C'C)^{-1}C'$ and $M[C] = I - P[C]$. If $C = [A \,\vdots\, B]$, we see on applying standard formulae for inverting partitioned matrices (see Schmidt (1976, Section 1.5, p. 27)) that $M[C] = M[B] - P[M[B]A]$, from which (2.9) follows.

where

$$\iota(\lambda)=\begin{bmatrix}1\\ \lambda\\ \lambda^2\\ \cdot\\ \cdot\\ \cdot\\ \lambda^{T-1}\end{bmatrix},\quad C(\lambda)=\begin{bmatrix}0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0\\ 1 & 0 & & & & & \cdot\\ \lambda & 1 & 0 & & & & \cdot\\ \lambda^2 & \cdot & \cdot & \cdot & & & \cdot\\ \cdot & & \cdot & \cdot & \cdot & & \cdot\\ \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot\\ \lambda^{T-2} & & \cdot & \cdot & \lambda & 1 & 0\end{bmatrix},$$

so that under $H(\lambda_0, \gamma_0)$,

$$\text{(2.11)}\qquad \mathbf{y}_{-1}/\sigma = y_0\iota(\lambda_0)/\sigma + C(\lambda_0)Z\gamma_0/\sigma + C(\lambda_0)\bar{Z}\bar{\gamma}/\sigma + C(\lambda_0)\varepsilon/\sigma,$$

where σ and $\bar{\gamma}$ are unknown. Thus, $M[X]\mathbf{y}_{-1}/\sigma$ consists of four terms, among which only the fourth one is pivotal: in the first two terms the factor $1/\sigma$ causes problems, while $\bar{\gamma}/\sigma$ is a nuisance factor in the third one.

These nuisance terms can be removed by extending the regression. Extending a model by including particular redundant regressors in order to achieve invariance of tests has also been suggested in Dufour and Dagenais (1985) and Dufour and King (1991, p. 125) with respect to inference problems in model (1.1). For model (1.2) the basic idea originates from Kiviet and Phillips (1990, 1992). To get a pivotal statistic for testing the restrictions $\lambda = \lambda_0$ and $\gamma = \gamma_0$, rewrite model (2.4) as

$$\text{(2.12)}\qquad \begin{aligned}\mathbf{y} &= \lambda\big[\mathbf{y}_{-1} - y_0\iota(\lambda_0) - C(\lambda_0)Z\gamma_0 - C(\lambda_0)\bar{Z}\bar{\gamma}\big]\\ &\quad + \lambda\big[y_0\iota(\lambda_0) + C(\lambda_0)Z\gamma_0 + C(\lambda_0)\bar{Z}\bar{\gamma}\big] + Z\gamma + \bar{Z}\bar{\gamma} + \varepsilon\end{aligned}$$

and then consider a more general model by relaxing some of the coefficient restrictions in (2.12):

$$\text{(2.13)}\qquad \begin{aligned}\mathbf{y} &= \lambda\mathbf{y}_{-1} + Z\gamma + \bar{Z}\bar{\gamma} + \lambda_{**}\iota(\lambda_0)\\ &\quad + \dot{\lambda}_{**}[C(\lambda_0)Z\gamma_0] + C(\lambda_0)\bar{Z}\bar{\gamma}_{**} + \varepsilon,\end{aligned}$$

which specializes to model (2.4) when $\lambda_{**} = 0$, $\dot{\lambda}_{**} = 0$, and $\bar{\gamma}_{**} = 0$. We now consider testing the null hypothesis

$$\text{(2.14)}\qquad H^*(\lambda_0, \gamma_0): \lambda = \lambda_0, \gamma = \gamma_0,$$

against (2.13). One obvious statistic to do this is the LR-type statistic

$$\text{(2.15)}\qquad \begin{aligned}\mathscr{L}^*_{\lambda,\gamma}(\lambda_0, \gamma_0) &= (\mathbf{y} - \lambda_0\mathbf{y}_{-1} - Z\gamma_0)'M\big[\bar{Z}(\lambda_0, \gamma_0)\big]\\ &\quad \times(\mathbf{y} - \lambda_0\mathbf{y}_{-1} - Z\gamma_0)/\mathbf{y}'M[W(\lambda_0, \gamma_0)]\mathbf{y},\end{aligned}$$

where $\bar{Z}(\lambda_0, \gamma_0) = [\bar{Z} \mid \iota(\lambda_0) \mid C(\lambda_0)Z\gamma_0 \mid C(\lambda_0)\bar{Z}]$ and $W(\lambda_0, \gamma_0) = [\mathbf{y}_{-1} \mid Z \mid \bar{Z}(\lambda_0, \gamma_0)]$. An intuitive way to see that $\mathscr{L}^*_{\lambda,\gamma}(\lambda_0, \gamma_0)$ is pivotal under $H^*(\lambda_0, \gamma_0)$ consists in noting that the residuals $\hat{\mathbf{u}}_{-1} \equiv M[Z \mid \bar{Z}(\lambda_0, \gamma_0)]\mathbf{y}_{-1}$ of the regression of $\mathbf{y}_{-1}$ on $[Z \mid \bar{Z}(\lambda_0, \gamma_0)]$ do not involve β as a nuisance parameter, and since the residuals of the regression of $\mathbf{y}$ on $[\mathbf{y}_{-1} \vdots Z \vdots \bar{Z}(\lambda_0, \gamma_0)]$ satisfy $M[\mathbf{y}_{-1} \vdots Z \vdots \bar{Z}(\lambda_0, \gamma_0)]\mathbf{y} = M[\mathbf{y}_{-1} \vdots Z \vdots \bar{Z}(\lambda_0, \gamma_0)]\varepsilon = M[\hat{\mathbf{u}}_{-1} \vdots Z \vdots \bar{Z}(\lambda_0, \gamma_0)]\varepsilon$, the distribution of the residual sum of squares $\mathbf{y}'M[\mathbf{y}_{-1} \vdots Z \vdots \bar{Z}(\lambda_0, \gamma_0)]\mathbf{y} = \varepsilon' M[\hat{\mathbf{u}}_{-1} \vdots Z \vdots \bar{Z}(\lambda_0, \gamma_0)]\varepsilon$ does not depend on β. By a similar argument, the distribution of the numerator of $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0, \gamma_0)$ does not depend on β, and taking the ratio of both sums of squares then eliminates σ from the null distribution of $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0, \gamma_0)$. Proposition 1 below provides a more rigorous proof of the pivotal character of $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0, \gamma_0)$ under $H^{*}(\lambda_0, \gamma_0)$.

It is clear that test statistic (2.15) does not take into account a few valid restrictions. Therefore, we shall also consider testing

$$(2.16) \qquad H^{**}(\lambda_0, \gamma_0)\colon \lambda = \lambda_0,\ \gamma = \gamma_0,\ \lambda_{**} = 0,\ \dot{\lambda}_{**} = 0,\ \bar{\gamma}_{**} = 0$$

in model (2.13) with the test statistic

$$(2.17) \qquad \mathscr{L}^{**}_{\lambda,\gamma}(\lambda_0, \gamma_0) = (\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)' \times M[\bar{Z}](\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)/\mathbf{y}'M[W(\lambda_0, \gamma_0)]\mathbf{y}.$$

When y_0 is taken as fixed, we can also obtain pivotal statistics from extended regressions with one redundant regressor less. To that end we consider the hypotheses

$$(2.18) \qquad H^{\dagger}(\lambda_0, \gamma_0)\colon \lambda = \lambda_0,\ \gamma = \gamma_0,$$

$$(2.19) \qquad H^{\dagger *}(\lambda_0, \gamma_0)\colon \lambda = \lambda_0,\ \gamma = \gamma_0,\ \lambda_{*} = 0,\ \bar{\gamma}_{**} = 0,$$

in the extended model

$$(2.20) \qquad \mathbf{y} = \lambda \mathbf{y}_{-1} + Z\gamma + \bar{Z}\bar{\gamma} + \lambda_{*}[y_0 \iota(\lambda_0) + C(\lambda_0)Z\gamma_0] + C(\lambda_0)\bar{Z}\bar{\gamma}_{**} + \varepsilon,$$

which is another generalization of (2.12). We test $H^{\dagger}(\lambda_0, \gamma_0)$ using

$$(2.21) \qquad \mathscr{L}^{\dagger}_{\lambda,\gamma}(\lambda_0, \gamma_0) = (\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)' M\big[\bar{Z}^{\dagger}(\lambda_0, \gamma_0)\big] \times (\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)/\mathbf{y}'M\big[W^{\dagger}(\lambda_0, \gamma_0)\big]\mathbf{y},$$

where $\bar{Z}^{\dagger}(\lambda_0, \gamma_0) = [\bar{Z} \vdots y_0 \iota(\lambda_0) + C(\lambda_0)Z\gamma_0 \vdots C(\lambda_0)\bar{Z}]$ and $W^{\dagger}(\lambda_0, \gamma_0) = [\mathbf{y}_{-1} \vdots Z \vdots \bar{Z}^{\dagger}(\lambda_0, \gamma_0)]$. Similarly, $H^{\dagger *}(\lambda_0, \gamma_0)$ is tested with

$$(2.22) \qquad \mathscr{L}^{\dagger *}_{\lambda,\gamma}(\lambda_0, \gamma_0) = (\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)' \times M[\bar{Z}](\mathbf{y} - \lambda_0 \mathbf{y}_{-1} - Z\gamma_0)/\mathbf{y}'M\big[W^{\dagger}(\lambda_0, \gamma_0)\big]\mathbf{y}.$$

The basic properties that enable one to use the test procedures suggested above for obtaining exact inference are stated in the following proposition.

Proposition 1: *Suppose that model* (2.4) *and Assumption A hold. Then the statistics* $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0, \gamma_0)$, $\mathscr{L}^{**}_{\lambda,\gamma}(\lambda_0, \gamma_0)$, $\mathscr{L}^{\dagger}_{\lambda,\gamma}(\lambda_0, \gamma_0)$, *and* $\mathscr{L}^{\dagger *}_{\lambda,\gamma}(\lambda_0, \gamma_0)$, *given in* (2.15), (2.17), (2.21), *and* (2.22) *respectively, can be written as follows when* $\lambda = \lambda_0$

and $\gamma = \gamma_0$:

$$\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0,\gamma_0) = \frac{\eta' M\left[\bar{Z}(\lambda_0,\gamma_0)\right]\eta}{S[\eta, Z(\lambda_0,\gamma_0)]},$$

$$\mathscr{L}^{**}_{\lambda,\gamma}(\lambda_0,\gamma_0) = \frac{\eta' M[\bar{Z}]\eta}{S[\eta, Z(\lambda_0,\gamma_0)]},$$

$$\mathscr{L}^{\dagger}_{\lambda,\gamma}(\lambda_0,\gamma_0) = \frac{\eta' M\left[\bar{Z}^{\dagger}(\lambda_0,\gamma_0)\right]\eta}{S\left[\eta, Z^{\dagger}(\lambda_0,\gamma_0)\right]},$$

$$\mathscr{L}^{\dagger *}_{\lambda,\gamma}(\lambda_0,\gamma_0) = \frac{\eta' M[\bar{Z}]\eta}{S\left[\nu, Z^{\dagger}(\lambda_0,\gamma_0)\right]},$$

where $\eta = \varepsilon/\sigma$ *and*

$$S(\eta,A) = \eta' M[A]\eta - \left[(\eta' M[A]C(\lambda_0)\eta)^2/\eta' C(\lambda_0)', M[A]C(\lambda_0)\eta\right], \quad \textit{for}$$

$$A = Z(\lambda_0,\gamma_0) = \left[Z : \bar{Z}(\lambda_0,\gamma_0)\right] = \left[Z : \bar{Z} : \iota(\lambda_0) : C(\lambda_0)Z\gamma_0 : C(\lambda_0)\bar{Z}\right] \quad \textit{or}$$

$$A = Z^{\dagger}(\lambda_0,\gamma_0) = \left[Z : \bar{Z}^{\dagger}(\lambda_0,\gamma_0)\right] = \left[Z : \bar{Z} : y_0\iota(\lambda_0) + C(\lambda_0)Z\gamma_0 : C(\lambda_0)\bar{Z}\right] \quad \textit{respectively}.$$

Proposition 1 shows that the null distributions of the test statistics considered do not involve the nuisance parameters $\bar{\gamma}$ and σ, as long as the distribution of ε/σ does not. The null distributions of $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ and $\mathscr{L}^{**}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ only depend on λ_0, γ_0, R, and the regressor matrix X, as well as on the distribution of ε/σ; those of $\mathscr{L}^{\dagger}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ and $\mathscr{L}^{\dagger *}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ also depend on y_0. For example, irrespective of the values of $\bar{\gamma}$ and σ, there is a point $\tilde{\mathscr{L}}^{*}_{\lambda,\gamma}(\alpha;\lambda_0,\gamma_0,X,R)$ such that $P[\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0,\gamma_0) \geq \tilde{\mathscr{L}}^{*}_{\lambda,\gamma}(\alpha;\lambda_0,\gamma_0,X,R) | \lambda = \lambda_0, \gamma = \gamma_0] = \alpha$. Thus the test that rejects the null when $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0,\gamma_0) > \tilde{\mathscr{L}}^{*}_{\lambda,\gamma}(\alpha:\lambda_0,\gamma_0,X,R)$ has size α. An exact confidence set for (λ,γ) with size $1-\alpha$ is given by

$$\mathscr{C}^{*}_{\lambda,\gamma}(\alpha) = \left\{(\lambda_0,\gamma_0) : \mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0,\gamma_0) \leq \tilde{\mathscr{L}}^{*}_{\lambda,\gamma}(\alpha;\lambda_0,\gamma_0,X,R)\right\}. \tag{2.23}$$

However, the critical points $\tilde{\mathscr{L}}^{*}_{\lambda,\gamma}(\alpha;\lambda_0,\gamma_0,X,R)$ are difficult to evaluate with great precision. In Section 5 we show how closely related fully exact tests and confidence sets can be obtained from a finite and relatively small number of Monte Carlo experiments.

When $r = 0$, the above statistics provide tests of $\lambda = \lambda_0$. In particular, the statistic (2.15) can be expressed as

$$\mathscr{L}^{*}_{\lambda}(\lambda_0) = (\mathbf{y} - \lambda_0\mathbf{y}_{-1})' M[X(\lambda_0)](\mathbf{y}_0 - \lambda\mathbf{y}_{-1})/\mathbf{y}' M[W(\lambda_0)]\mathbf{y}, \tag{2.24}$$

where $X(\lambda_0) = [X \,\vdots\, \iota(\lambda_0) \,\vdots\, C(\lambda_0)X]$ and $W(\lambda_0) = [\mathbf{y}_{-1} \,\vdots\, X(\lambda_0)]$, while the statistic in (2.17) can be written

$$\mathscr{L}_\lambda^{**}(\lambda_0) = (\mathbf{y} - \lambda_0\mathbf{y}_{-1})'M[X](\mathbf{y} - \lambda_0\mathbf{y}_{-1})/\mathbf{y}'M[W(\lambda_0)]\mathbf{y}. \tag{2.25}$$

If $y_0 = 0$, (2.24) and (2.25) specialize respectively to

$$\mathscr{L}_\lambda^{\dagger}(\lambda_0) = (\mathbf{y} - \lambda_0\mathbf{y}_{-1})'M\left[X^{\dagger}(\lambda_0)\right](\mathbf{y} - \lambda_0\mathbf{y}_{-1})/\mathbf{y}'M\left[W^{\dagger}(\lambda_0)\right]\mathbf{y} \tag{2.26}$$

and

$$\mathscr{L}_\lambda^{\dagger *}(\lambda_0) = (\mathbf{y} - \lambda_0\mathbf{y}_{-1})'M[X](\mathbf{y} - \lambda_0\mathbf{y}_{-1})/\mathbf{y}'M\left[W^{\dagger}(\lambda_0)\right]\mathbf{y}, \tag{2.27}$$

where $X^{\dagger}(\lambda_0) = [X \,\vdots\, C(\lambda_0)X]$ and $W^{\dagger}(\lambda_0) = [\mathbf{y}_{-1} \,\vdots\, X^{\dagger}(\lambda_0)]$.

The $\mathscr{L}_\lambda^{**}(\lambda_0)$ procedure is a generalization of the test denoted Φ_3 in Dickey and Fuller (1981, p. 1059) for the special case where $\lambda_0 \equiv 1$ and $X \equiv \iota(1)$. Also note that the tests $\mathscr{L}_\lambda^{*}(\lambda_0)$ and $\mathscr{L}_\lambda^{\dagger}(\lambda_0)$ restrict only one parameter, and therefore testing against one-sided alternatives is possible here. The t ratio version of $\mathscr{L}_\lambda^{*}(\lambda_0)$ was already suggested in Kiviet and Phillips (1990). Finally, we note that

$$\mathscr{C}_\lambda^{*}(\alpha) = \left\{\lambda_0 \in \mathscr{D}_\lambda : \mathscr{L}_\lambda^{*}(\lambda_0) \le \tilde{\mathscr{L}}_\lambda^{*}(\alpha;\lambda_0, X)\right\}, \tag{2.28}$$

where $P[\mathscr{L}_\lambda^{*}(\lambda_0) > \tilde{\mathscr{L}}_\lambda^{*}(\alpha;\lambda_0, X)|\lambda = \lambda_0] = \alpha$, is a confidence set for λ with level $1-\alpha$. Confidence sets based on the other three tests for inference on λ are built in the same way.

3. INFERENCE ON RESTRICTIONS NOT INVOLVING λ EXPLICITLY

Joint inference on λ and any linear transformation of β can also be obtained from a two-stage procedure, i.e., an exact simultaneous confidence set for (λ, γ) can be constructed by combining an exact confidence set for λ with the corresponding family of "conditional" confidence sets for γ; see Dufour (1990). The duality between tests and confidence sets then again leads to an exact test for any joint null hypothesis on λ and γ. We already proposed a procedure for such joint tests in Section 2, but the two-stage procedure makes it easier to obtain exact inference on restrictions not involving λ.

First assume that the true value of λ is given. An exact similar test for $\gamma = \gamma_0$ is then obtained (without extending the model) from the statistic

$$\begin{aligned}\mathscr{L}_{\gamma|\lambda}(\gamma_0) = {} & (\mathbf{y} - \lambda\mathbf{y}_{-1} - Z\gamma_0)'M[\bar{Z}] \\ & \times(\mathbf{y} - \lambda\mathbf{y}_{-1} - Z\gamma_0)/(\mathbf{y} - \lambda\mathbf{y}_{-1})'M[X](\mathbf{y} - \lambda\mathbf{y}_{-1}),\end{aligned} \tag{3.1}$$

which reduces when $\gamma = \gamma_0$ to $\mathscr{L}_{\gamma|\lambda}(\gamma_0) = \varepsilon'M[\bar{Z}]\varepsilon/\varepsilon'M[X]\varepsilon$. Given Assumption B, it yields a similar test, and its critical value $\tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha; X, R)$ is the smallest point such that $P[\mathscr{L}_{\gamma|\lambda}(\gamma_0) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha; X, R)] \le \alpha$; this critical value does not depend on γ_0 or λ. Then the set

$$\mathscr{C}_{\gamma|\lambda}(\alpha) = \left\{\gamma_0 : \mathscr{L}_{\gamma|\lambda}(\gamma_0) \le \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha; X, R)\right\} \tag{3.2}$$

is an exact confidence set for γ (given λ). Note that the statistic

$$(3.3)\qquad \mathscr{F}_{\gamma|\lambda}(\gamma_0) = \frac{T-k}{r}\left[\mathscr{L}_{\gamma|\lambda}(\gamma_0) - 1\right]$$

produces the same inferences. When $\varepsilon/\sigma \sim N[0, I_T]$, $\mathscr{F}_{\gamma|\lambda}(\gamma_0)$ follows (under the null hypothesis) a Fisher distribution with $(r, T-k)$ degrees of freedom. So, under normality, $\tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha; X, R) = 1 + r(T-k)^{-1}\tilde{\mathscr{F}}(\alpha; r, T-k)$.

Suppose now that λ is unknown, and let

$$(3.4)\qquad 0 < \alpha = \alpha_1 + \alpha_2 < 1 \quad \text{with} \quad 0 \le \alpha_i < 1 \qquad (i = 1,2).$$

A size $1 - \alpha_1$ confidence set $\mathscr{C}_\lambda(\alpha_1)$ for λ can be constructed by using $\mathscr{L}_\lambda^*(\lambda_0)$ or $\mathscr{L}_\lambda^{**}(\lambda_0)$ as described in Section 2. Now consider the set

$$(3.5)\qquad \mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2) = \left\{(\lambda_0, \gamma_0) : \lambda_0 \in \mathscr{C}_\lambda(\alpha_1) \text{ and } \gamma_0 \in \mathscr{C}_{\gamma|\lambda_0}(\alpha_2)\right\},$$

where $\mathscr{C}_{\gamma|\lambda_0}(\alpha_2) = \{\gamma_0 : \mathscr{L}_{\gamma|\lambda_0}(\lambda_0, \gamma_0) \le \tilde{\mathscr{L}}_{\gamma|\lambda_0}(\alpha_2; X, R)\}$. Under $(\lambda, \gamma) = (\lambda_0, \gamma_0)$, we have $\mathscr{L}_{\gamma|\lambda}(\gamma_0) = \varepsilon' M[\bar{Z}]\varepsilon / \varepsilon' M[X]\varepsilon$, and so the quantiles of $\mathscr{L}_{\gamma|\lambda}(\gamma_0)$ do not depend on nuisance parameters nor on λ_0 and γ_0. This invariance is especially convenient when $\mathscr{C}_{\gamma|\lambda_0}(\alpha_2)$ has to be assessed for various λ_0 values, which—as we shall see—is required when these conditional sets are used to construct unconditional inference for (λ, γ) or γ only.

Upon using the Boole-Bonferroni inequality, we find

$$(3.6)\qquad \begin{aligned} &P\left[(\lambda, \gamma) \in \mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2)\right] \\ &= P\left[\lambda \in \mathscr{C}_\lambda(\alpha_1) \text{ and } \gamma \in \mathscr{C}_{\gamma|\lambda}(\alpha_2)\right] \\ &\ge 1 - P[\lambda \notin \mathscr{C}_\lambda(\alpha_1)] - P\left[\gamma \notin \mathscr{C}_{\gamma|\lambda}(\alpha_2)\right] \\ &= 1 - \alpha_1 - \alpha_2 = 1 - \alpha, \end{aligned}$$

so that $\mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2)$ is a conservative $1 - \alpha$ level confidence set. This two-stage joint confidence set will be quite convenient for making inference on individual elements of β and cross-restrictions between λ and β.

3.1. *Inference on β Only*

Below we shall "marginalize" with respect to λ the two-stage joint confidence set $\mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2)$ defined in (3.5) in order to produce inference about γ only. To do this, we define the r-dimensional random sets

$$(3.7)\qquad \begin{aligned} \mathscr{C}_\gamma^U(\alpha_1, \alpha_2) &= \left\{\gamma_0 : \exists \lambda_0 \in \mathscr{C}_\lambda(\alpha_1) \text{ such that } (\lambda_0, \gamma_0) \in \mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2)\right\}, \\ \mathscr{C}_\gamma^L(\alpha_1, \alpha_2) &= \left\{\gamma_0 : \forall \lambda_0 \in \mathscr{C}_\lambda(\alpha_1), (\lambda_0, \gamma_0) \in \mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2)\right\}, \end{aligned}$$

with the convention $\mathscr{C}_\gamma^U(\alpha_1, \alpha_2) = \mathscr{C}_\gamma^L(\alpha_1, \alpha_2) = \varnothing$ when $\mathscr{C}_\lambda(\alpha_1) = \varnothing$, where $\varnothing$ is the empty set. The two sets are nested, i.e., $\varnothing \subseteq \mathscr{C}_\gamma^L(\alpha_1, \alpha_2) \subseteq \mathscr{C}_\gamma^U(\alpha_1, \alpha_2)$, and their use is justified by the following proposition.

PROPOSITION 2: *Suppose that model* (2.4) *holds jointly with the Assumptions A and B. Let* $\mathscr{C}_\lambda(\alpha_1)$ *be a confidence set for* λ *such that* $P[\lambda \in \mathscr{C}_\lambda(\alpha_1)] \geq 1 - \alpha_1$. *If* $0 \leq \alpha_1 < 1$, $0 < \alpha_1 + \alpha_2 = \alpha \leq 1 - \alpha_1$ *and* $\alpha_2' = \alpha + \alpha_1 < 1$, *then the random sets defined in* (3.7) *satisfy* $\mathscr{C}_\gamma^L(\alpha_1, \alpha_2') \subseteq \mathscr{C}_\gamma^U(\alpha_1, \alpha_2)$ *and*

$$P\left[\gamma \in \mathscr{C}_\gamma^L(\alpha_1, \alpha_2')\right] \leq 1 - \alpha \leq P\left[\gamma \in \mathscr{C}_\gamma^U(\alpha_1, \alpha_2)\right].$$

From the above results, conservative and liberal tests for $\gamma = \gamma_0$ can be derived. We define

$$(3.8) \qquad \mathscr{L}_\gamma^L(\gamma_0) = \inf\left\{\mathscr{L}_{\gamma|\lambda_0}(\gamma_0) : \lambda_0 \in \mathscr{C}_\lambda(\alpha_1)\right\},$$
$$\mathscr{L}_\gamma^U(\gamma_0) = \sup\left\{\mathscr{L}_{\gamma|\lambda_0}(\gamma_0) : \lambda_0 \in \mathscr{C}_\lambda(\alpha_1)\right\}.$$

Note the event $\gamma_0 \in \mathscr{C}_\gamma^U(\alpha_1, \alpha_2)$ implies $\mathscr{L}_\gamma^L(\gamma_0) \leq \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2; X, R)$. With respect to the true γ this entails for $\gamma = \gamma_0$:

$$(3.9) \qquad P\left[\mathscr{L}_\gamma^L(\gamma_0) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2; X, R)\right] \leq P\left[\gamma \notin \mathscr{C}_\gamma^U(\alpha_1, \alpha_2)\right] \leq \alpha.$$

Similarly, $\mathscr{L}_\gamma^U(\gamma_0) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R)$ is equivalent with $\gamma_0 \notin \mathscr{C}_\gamma^L(\alpha_1, \alpha_2')$, hence

$$(3.10) \qquad P\left[\mathscr{L}_\gamma^U(\gamma_0) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R)\right] = P\left[\gamma \notin \mathscr{C}_\gamma^L(\alpha_1, \alpha_2')\right] \geq \alpha,$$

when $\gamma = \gamma_0$. Because $\alpha_2' \geq \alpha_2$ entails $\tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2; X, R) \geq \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R)$, the generalized bounds principle can be applied, yielding the test:

$$(3.11) \qquad \text{reject } \gamma = \gamma_0 \text{ when } \mathscr{L}_\gamma^L(\gamma_0) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2; X, R),$$
$$\text{accept } \gamma = \gamma_0 \text{ when } \mathscr{L}_\gamma^U(\gamma_0) \leq \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R),$$

whereas the test is inconclusive otherwise. In practice, simulation procedures have to be used first to obtain a $\mathscr{C}_\lambda(\alpha_1)$ interval, and then $\mathscr{L}_\gamma^L(\gamma_0)$ and $\mathscr{L}_\gamma^U(\gamma_0)$ have to be assessed. Under normality of ε the $\tilde{\mathscr{L}}_{\gamma|\lambda}$ can easily be obtained from tabulated critical F values. In general, they have to be simulated. However, in Section 5 we will make this procedure operational in an easier way.

3.2. *Cross-restrictions between* λ *and* β

As far as linear restrictions on λ and β are concerned, the only situation for which we have not yet obtained an exact inference procedure is the case where the restrictions involve both λ and elements of β, but do not explicitly (nor implicitly) specify $\lambda = \lambda_0$. Let us focus on the case

$$(3.12) \qquad H_0(\kappa_0) : \lambda + \kappa_1' \beta = \kappa_0,$$

where only one restriction is tested; κ_1 is a $k \times 1$ vector and κ_0 a scalar, both of which are known. This case can be handled as follows. Let

$$(3.13) \qquad \gamma = \kappa_1' \beta, \quad \kappa = \lambda + \gamma = \lambda + \kappa_1' \beta.$$

Hence, both γ and κ are scalar unknown parameters here. We define the set

$$(3.14) \qquad \mathscr{C}_{\kappa}(\alpha) = \left\{\kappa_0 : \exists(\lambda_0, \gamma_0) \in \mathscr{C}_{\lambda,\gamma}(\alpha) \text{ such that } \kappa_0 = \lambda_0 + \gamma_0\right\},$$

where $\mathscr{C}_{\lambda,\gamma}(\alpha)$ is a level $1-\alpha$ confidence set for (λ, γ). Then we have

$$(3.15) \qquad P[\kappa \in \mathscr{C}_{\kappa}(\alpha)] \geq P\left[(\lambda, \gamma) \in \mathscr{C}_{\lambda,\gamma}(\alpha)\right] \geq 1-\alpha,$$

and the test which rejects $H_0(\kappa_0): \kappa = \kappa_0$, when $\kappa_0 \notin \mathscr{C}_{\kappa}(\alpha)$ has level α. The case where we have more general restrictions, including at least one of the form (3.12), is covered by the general results in the next section.

4. TESTS OF NON-REGULAR HYPOTHESES

We now generalize the foregoing results on linear restrictions and consider general problems where two arbitrary hypotheses are compared, viz.

$$(4.1) \qquad h_0(\lambda, \gamma) = 0 \Leftrightarrow (\lambda, \gamma')' \in \Gamma_0 \subseteq \mathbb{R}^{r+1}, \qquad \Gamma_0 \neq \varnothing,$$

$$(4.2) \qquad h_1(\lambda, \beta) = 0 \Leftrightarrow (\lambda, \beta')' \in \Omega_1 \subseteq \mathbb{R}^{k+1}, \qquad \Omega_1 \neq \varnothing.$$

These cover special cases such as nonlinear restrictions, inequality restrictions, and non-nested hypotheses.

When λ is known, $\mathbf{y} - \lambda \mathbf{y}_{-1} = \mathbf{y}(\lambda) = X\beta + \varepsilon$ satisfies the assumptions of the standard linear regression model, except that ε/σ may be non-normal. Dufour (1989) deals with exact solutions to inference problems on nonregular hypotheses in that model under normality. We shall stick to our more general Assumptions A and B and extend these results to dynamic models. Let

$$(4.3) \qquad S(\lambda_0, \beta_0) = [\mathbf{y}(\lambda_0) - X\beta_0]'[\mathbf{y}(\lambda_0) - X\beta_0]$$

denote the residual sum of squares for arbitrary coefficient values $(\lambda_0, \beta_0')'$, where $(\lambda_0, \beta_0')' \in \mathbb{R}^{k+1}$ with $\lambda_0 \in \mathscr{D}_{\lambda}$. We first consider the case where the true value of λ is known, and define

$$(4.4) \qquad S_0(\lambda) = \inf\{S(\lambda, \beta_0) : \beta_0 \in \mathbb{R}^k \text{ and } h_0(\lambda, R\beta_0) = 0\},$$

$$S_1(\lambda) = \inf\{S(\lambda, \beta_0) : \beta_0 \in \mathbb{R}^k \text{ and } h_1(\lambda, \beta_0) = 0\}.$$

The LR type test for the two hypotheses can be written (after a monotonic transformation, and still taking λ to be known) as:

$$(4.5) \qquad \mathscr{L}_{\Gamma|\lambda}(\Gamma_0, \Omega_1) = S_0(\lambda)/S_1(\lambda).$$

We also define for any given value γ_0 the sum of squares:

$$(4.6) \qquad \bar{S}_0(\lambda, \gamma_0) = \inf\{S(\lambda, \beta_0) : \beta_0 \in \mathbb{R}^k \text{ and } R\beta_0 = \lambda_0 \in \mathbb{R}^r\},$$

$$(4.7) \qquad S(\lambda) = \inf\{S(\lambda, \beta_0) : \beta_0 \in \mathbb{R}^k\}.$$

Note that the statistic (3.1) for testing $\gamma = \gamma_0$ can now be expressed as $\mathscr{L}_{\gamma|\lambda}(\gamma_0) = \bar{S}_0(\lambda, \gamma_0)/S(\lambda)$. Below it is shown that its critical value $\tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha; X, R)$ remains valid as a conservative critical value for testing the nonregular hypotheses (4.1) and (4.2) through the statistic $\mathscr{L}_{\Gamma|\lambda}(\Gamma_0, \Omega_1)$.

PROPOSITION 3: *Suppose that model* (2.4) *holds jointly with the Assumptions A and B. Then, if* $(\lambda, \gamma')' \in \Gamma_0$, *we have* $P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0, \Omega_1) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha; X, R)] \leq \alpha$.

It is also possible to obtain liberal critical values, but these are of less interest. Proposition 3 concerns the unrealistic case where λ is known. However, an exact confidence interval $\mathscr{C}_\lambda(\alpha)$ can be exploited to obtain unconditional inference on the hypotheses (4.1) and (4.2). We define

$$(4.8)\qquad \mathscr{L}_\Gamma^L(\Gamma_0, \Omega_1) = \inf\left\{\mathscr{L}_{\Gamma|\lambda_0}(\Gamma_0, \Omega_1) : \lambda_0 \in \mathscr{C}_\lambda(\alpha_1)\right\},$$
$$\mathscr{L}_\Gamma^U(\Gamma_0, \Omega_1) = \sup\left\{\mathscr{L}_{\Gamma|\lambda_0}(\Gamma_0, \Omega_1) : \lambda_0 \in \mathscr{C}_\lambda(\alpha_1)\right\}.$$

PROPOSITION 4: *Under the assumptions of Proposition* 3, *let* $\mathscr{C}_\lambda(\alpha_1)$ *be a confidence set for* λ *such that* $P[\lambda \in \mathscr{C}_\lambda(\alpha_1)] \geq 1 - \alpha_1$. *If* $0 \leq \alpha_1 < 1$, $0 < \alpha_1 + \alpha_2 = \alpha \leq 1 - \alpha_1$, *and* $\alpha_2' = \alpha + \alpha_1$, *then for the statistics defined in* (4.8) *we have under* $(\lambda, \gamma')' \in \Gamma_0$:

$$P\left[\mathscr{L}_\Gamma^L(\Gamma_0, \Omega_1) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2; X, R)\right]$$
$$\leq \alpha \leq P\left[\mathscr{L}_\Gamma^U(\Gamma_0, \Omega_1) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R)\right].$$

This yields the following generalized bounds test procedure:

$$(4.9)\qquad \text{reject}(\lambda, \gamma')' \in \Gamma_0 \text{ when } \mathscr{L}_\Gamma^L(\Gamma_0, \Omega_1) > \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2; X, R),$$
$$\text{accept } (\lambda, \gamma')' \in \Gamma_0 \text{ when } \mathscr{L}_\Gamma^U(\Gamma_0, \Omega_1) \leq \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R),$$

whereas the test is inconclusive otherwise. Rather than obtaining a test for $(\lambda, \gamma')' \in \Gamma_0$ from the above procedure, which is inspired by the two-stage approach of Section 3, we can also make use of the comprehensive procedures of Section 2. We shall illustrate this for a particular example.

In the context of dynamic model (1.2), which originates from the underlying relationship (2.1), nonlinear transformations of the regression coefficients which are particularly relevant are expressions such as: $(\delta_{j0} + \cdots + \delta_{jL(j)})/(1 - \lambda)$, i.e. the total multiplier for regressor $z^{(j)}$, where $j = 1, \ldots, J; 1 - (\lambda)^{i+1}$, i.e. the standardized interim multiplier after i time periods; $\lambda/(1 - \lambda)$, the mean lag; and $\max\{\ln(0.5)/\ln(\lambda) - 1, 0\}$, i.e. the median lag. Note that these nonlinear characterizations of particular aspects of the dynamic adjustment process are only meaningful if $\mathscr{D}_\lambda \subseteq (-1, +1)$.

We shall focus now on the problem of producing exact inference on the parameter $\varphi = \gamma/(1 - \lambda)$, where again $\gamma = R\beta$; hence, φ may represent a vector of total or long-run multipliers. From a confidence set $\mathscr{C}_{\lambda,\gamma}(\alpha)$ for (λ, γ), we construct the set

$$(4.10)\qquad \mathscr{C}_\varphi(\alpha) = \left\{\varphi_0 : \exists(\lambda_0, \gamma_0) \in \mathscr{C}_{\lambda,\gamma}(\alpha) \text{ such that } \varphi_0 = \gamma_0/(1 - \lambda_0)\right\}.$$

Note that this set is not necessarily bounded. It is clear that $P[\varphi \in \mathscr{C}_\varphi(\alpha)] \geq 1 - \alpha$, and that an exact test of $\varphi = \varphi_0$ corresponds to this confidence set. In case we start from the two-state confidence set $\mathscr{C}_{\lambda,\gamma}(\alpha_1, \alpha_2)$ in (3.6) we can obtain the set, which corresponds to test procedure (4.9):

$$(4.11) \qquad \mathscr{C}_\varphi(\alpha_1, \alpha_2) = \Big\{ \varphi_0 : \varphi_0 = \gamma_0/(1 - \lambda_0) \text{ and } \hat{\mathscr{L}}_{\gamma|\lambda_0}(\gamma_0, \lambda_0) \leq \tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2'; X, R) \text{ for some } \lambda_0 \in \mathscr{C}_\lambda(\alpha_1) \Big\}.$$

5. EXACT INFERENCE BASED ON FEW MONTE CARLO EXPERIMENTS

When we want to use the various tests mentioned in Proposition 1 in practice, we face particular computational problems. The explicit assessment of critical values requires iterations, but that can be avoided by employing p values, both for testing and for the construction of confidence sets. We show this for the test based on $\mathscr{L}^*_{\lambda,\gamma}(\lambda_0, \gamma_0)$. Let $\hat{\mathscr{L}}^*_{\lambda,\gamma}(\lambda_0, \gamma_0)$ denote the value of this test statistic as obtained from the sample (it is simply a ratio of two particular residual sums of squares). Under the restrictions, the test statistic can be written as

$$(5.1) \qquad \eta' M\big[\bar{Z}(\lambda_0, \gamma_0)\big]\eta \times \left[\eta' M[Z(\lambda_0, \gamma_0)]\eta - \frac{(\eta' M[Z(\lambda_0, \gamma_0)]C(\lambda_0)\eta)^2}{\eta' C(\lambda_0)' M[Z(\lambda_0, \gamma_0)]C(\lambda_0)\eta} \right]^{-1}.$$

Drawings distributed as (5.1) can be obtained easily by computer simulation. When generating $N-1$ independent realizations of (5.1), indicated by $\{\mathscr{L}_j;\ j = 1, \ldots, N-1\}$, the frequency of the occurrence of $\mathscr{L}_j > \hat{\mathscr{L}}^*_{\lambda,\gamma}(\lambda_0, \gamma_0)$ gives an estimate of the p value. The restrictions are rejected if this is strictly smaller than α. Then (λ_0, γ_0) is not in the confidence set (2.23). By searching over $\mathscr{D}_\lambda$ and feasible γ_0 values the full set (2.23) can be established to a certain degree of precision, depending on the number N and the intricacy of the (grid-) search over the (λ, γ) space. As it happens, the number of replications has in fact to be extremely high. Using a normal approximation, the standard deviation of the p value estimate in the neighborhood of α is $[\alpha(1-\alpha)/N]^{1/2}$. Keeping this below 1% of α requires $N > 10^4(1-\alpha)/\alpha$; i.e. at $\alpha = 0.05$ we need $N \geq 190{,}000$ which seems prohibitively large.

To avoid this difficulty, we shall use instead "randomized" or "Monte Carlo" versions of the tests of Proposition 1, which yield genuinely exact and computationally much cheaper procedures. This approach was originally suggested by Dwass (1957), Barnard (1963), and Birnbaum (1974). It is based on a general property which we state for convenience in the following lemma.

LEMMA: *Let Z_j, $j = 1, \ldots, N$, be independent and identically distributed (i.i.d.) real random variables with continuous distribution, and let $R_j(N) = \sum_{i=1}^{N} U(Z_j - Z_i)$ be the rank of Z_j when $Z_1, \ldots, Z_N$ are ranked in nondecreasing order ($j = 1, \ldots, N$),* *where* $U(x)=1$ *if* $x \geq 0$, *and* $U(x)=0$ *if* $x<0$. *Then, for* $j=1,\dots,N$,

$$(5.2) \qquad P\left[R_j(N)/N \geq x\right] = (1+I[N(1-x)])/N, \quad \text{if} \quad 0<x\leq 1$$

where $I[x]$ *is the largest integer less than or equal to* x.

For a simple proof of this lemma, see Dufour and Kiviet (1994) or Hájek and Šidák (1967). Note that the lemma would not hold if $Z_1,\dots,Z_N$ were i.i.d. with a *discrete* distribution. We can use it as follows. Let Z_1 be the value of a test statistic computed from an observed sample, and $Z_2,\dots,Z_N$ i.i.d. random variables with the same distribution as Z_1 under a null hypothesis of interest. For example, to test $\lambda=\lambda_0$ and $\gamma=\gamma_0$, we may take $Z_1=\mathscr{L}_1^*(\lambda_0,\gamma_0)=\hat{\mathscr{L}}^*_{\lambda,\gamma}(\lambda_0,\gamma_0)$, the observed test statistic $\mathscr{L}^*_{\lambda,\gamma}(\lambda_0,\gamma_0)$, and $Z_j=\mathscr{L}_j^*(\lambda_0,\gamma_0)$, $j=2,\dots,N$, which are $N-1$ independent replications of the variable defined in (5.1). Clearly $Z_2,\dots,Z_N$ can easily be generated by Monte Carlo methods. Then $p_N=[N+1-R_1(N)]/N$ is (with probability one) the proportion of the variables $Z_1,Z_2,\dots,Z_N$ that are greater than or equal to Z_1, hence for $0<\alpha<1$ and provided $N\alpha$ is an integer, $P[p_N\leq\alpha]=P[R_1/N\geq(N+1)/N-\alpha]=\alpha$. p_N may thus be interpreted as a randomized p value, and can be used for building tests and confidence sets. The resulting procedures are not strictly speaking equivalent to the corresponding procedures one would obtain by computing analytically the appropriate critical values, but they nevertheless have the desired levels. In the simulation test an extra random element has been included; this yields genuine exactness for finite values of N. Given the fact that we cannot assess the relevant null distribution analytically, this simulation-based method is the only really exact way to proceed. There is typically a power loss associated with the simulation. However, as $N\to\infty$, the Monte Carlo procedure becomes equivalent to the corresponding nonrandomized procedure under weak regularity conditions; for further discussion, see Birnbaum (1974), Dwass (1957), Edgington (1980), Foutz (1980), and Jöckel (1986). For other applications of Monte Carlo tests in a time series context, see Dufour and Hallin (1987) and Theil and Shonkweiler (1986).

Since it is not clear which one of the various tests mentioned in Proposition 1 is the most powerful in a particular model, it is tempting (and perhaps sensible) to work with various procedures, and to compare their results before a choice is made. However, this will affect the significance level and so leaves room for further study of the relative power and interdependence of these individually exact tests.

Finally, it is worthwhile noting that so-called "parametric bootstrap tests" (see Efron and Tibshirani (1993, Section 6.5) or Hall (1992)) may be interpreted as Monte Carlo tests when the test statistic does not depend on nuisance parameters. In a Monte Carlo test the number of replications and the fact that we get a "randomized test" are explicitly taken into account (hence the possibility of getting an exact test), while in "bootstrap tests" the distribution of the test statistic is simply "approximated" by simulation and the number of replications is treated as being essentially infinite (which is of course never the case). So "parametric bootstrap tests" are not really alternative tests, but simply "approximate" Monte Carlo tests stripped of their finite sample justification. Because the literature on Monte Carlo tests considerably predates the "bootstrap" literature (which starts with Efron (1979) without reference to the earlier work of Dwass (1957), Barnard (1963), and Birnbaum (1974)), we prefer to use here the term "Monte Carlo test" rather than "bootstrap test." For further discussion of the relationship between Monte Carlo testing and the "bootstrap," see Hall and Titterington (1989) and Hall (1994). An examination of the qualities of the "nonparametric bootstrap" in autoregressive models is given in Giersbergen and Kiviet (1996).

6. SIMULATION RESULTS

Before we apply the procedures proposed above to empirical data, we will study their size and power properties relative to asymptotic tests in a few Monte Carlo experiments. We generated data according to the model:

$$(6.1)\qquad \begin{aligned} &y_t = \lambda y_{t-1} + \beta_1 + \beta_2 x_t + \beta_3 x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{IIN}(0, \sigma_\varepsilon^2),\\ &\Delta x_t = \delta + \nu_t, \qquad \nu_t \sim \text{IIN}(0, \sigma_\nu^2), \qquad y_{-S} = x_{-S} \equiv 0,\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad (t = -S, \ldots, T). \end{aligned}$$

In all experiments we set: $\sigma_\nu = 0.01$, $\delta = 0.01$, $\beta_1 = 0$, $\beta_2 = 1 - \lambda$, $\beta_3 = 0$, $S = 100$, $T = 20$. Hence, λ and σ_ε are the only parameters that will vary. Since ε_t and ν_t are mutually independent, x_t is strongly exogenous. For $|\lambda| < 1$, the variables y_t and x_t are cointegrated $I(1)$ processes. After substituting the selected parameter values, this partial adjustment model can be written in the error-correction form

$$(6.2)\qquad \Delta y_t = (1 - \lambda)\,\Delta x_t + (\lambda - 1)[y_{t-1} - x_{t-1}] + \varepsilon_t.$$

For $S \to \infty$, and using $[y_{t-1} - x_{t-1}] = -[\lambda\delta/(1 - \lambda)] + \Sigma_{i=0}^{\infty} \lambda^i(\varepsilon_{t-1-i} - \lambda\nu_{t-1-i})$, one finds

$$(6.3)\qquad \text{var}(\Delta y_t) = \left[(1 - \lambda)\sigma_\nu^2 + 2\sigma_\varepsilon^2\right]/(1 + \lambda).$$

The signal-to-noise ratio of (6.2) is given by $[\text{var}(\Delta y_t) - \text{var}(\varepsilon_t)]/\text{var}(\varepsilon_t) = (1 - \kappa)/\kappa$, where $\kappa = \text{var}(\varepsilon_t)/\text{var}(\Delta y_t) = (1 + \lambda)/[2 + (1 - \lambda)\sigma_\nu^2/\sigma_\varepsilon^2]$. For increasing σ_ε^2 the signal-to-noise ratio decreases, and κ converges to a maximum of $(1 + \lambda)/2$, which approaches unity as $\lambda \to 1$.

Table I contains results on tests for the hypothesis $\lambda = \lambda_0$ for a few positive λ_0 and λ values in combination with just two σ_ε^2 values, chosen such that κ is either $(1 + \lambda)/2 - 0.2$ or $(1 + \lambda)/2 - 10^{-4}$; the resulting R^2 values (averaged over the Monte Carlo replications) for the first equation of (6.1) are also given. We chose $N = 1000$ and generated 50,000 Monte Carlo replications. We kept x_t fixed and also the values of $\varepsilon_t/\sigma_\varepsilon$ for $t = -S, \ldots, 0$; other realizations of the x_t

TABLE I

SIZE (%) OF ASYMPTOTIC TESTS AND THE REJECTION FREQUENCIES (%) OF EXACT AND SIZE-CORRECTED ASYMPTOTIC TESTS IN MODEL (6.1); NOMINAL LEVEL $\alpha = 5\%$, $N = 1000$; 50,000 MONTE CARLO REPLICATIONS

λ_0	λ	κ	R^2	$\mathscr{F}$	$\mathscr{W}$	$t<$	$t>$
0.3	0.3	0.45	0.9860	6.1	11.6	10.8	1.0
0.3	0.3	0.6499	0.2084	6.2	11.9	11.5	0.6
0.5	0.5	0.55	0.9880	8.1	14.6	14.1	0.9
0.5	0.5	0.7499	0.3034	8.8	15.8	15.8	0.3
0.7	0.7	0.65	0.9919	9.3	16.1	15.6	0.7
0.7	0.7	0.8499	0.4729	13.8	23.5	23.8	0.2
0.9	0.9	0.75	0.9968	8.3	14.7	14.1	1.0
0.9	0.9	0.9499	0.6825	27.8	42.2	42.6	0.0

λ_0	λ	κ	R^2	$\mathscr{L}^*$	$\mathscr{L}^{**}$	$\mathscr{F}^C$	$\tilde{\mathscr{F}}^C$
0.3	0.3	0.45	0.9860	5.0	5.1	5.0	4.93
0.3	0.3	0.6499	0.2084	5.0	5.1	5.0	4.93
0.3	0.0	0.30	0.9866	22.4	15.0	29.8	4.93
0.3	0.0	0.4999	0.1684	22.2	12.7	25.1	4.93
0.3	0.6	0.60	0.9898	3.6	13.0	16.2	4.93
0.3	0.6	0.7999	0.3768	3.0	6.7	5.4	4.93
0.3	0.99	0.795	0.9988	100.0	100.0	100.0	4.93
0.3	0.99	0.9949	0.6962	19.8	26.7	35.0	4.93
0.5	0.5	0.55	0.9880	5.0	5.0	4.6	5.82
0.5	0.5	0.7499	0.3034	5.0	5.0	5.0	5.82
0.5	0.0	0.30	0.9866	39.9	29.4	57.5	5.82
0.5	0.0	0.4999	0.1684	38.1	22.7	47.4	5.82
0.5	0.99	0.795	0.9988	90.6	99.9	100.0	5.82
0.5	0.99	0.9949	0.6962	3.4	10.9	9.8	5.82
0.7	0.7	0.65	0.9919	5.1	4.9	3.4	7.23
0.7	0.7	0.8499	0.4729	5.1	4.9	5.0	7.23
0.7	0.99	0.795	0.9988	1.7	80.4	96.1	7.23
0.7	0.99	0.9949	0.6962	2.9	5.4	2.5	7.23
0.9	0.9	0.75	0.9968	5.1	4.9	1.1	10.16
0.9	0.9	0.9499	0.6825	5.1	4.9	5.0	10.16
0.9	0.99	0.795	0.9988	2.7	7.9	3.2	10.16
0.9	0.99	0.9949	0.6962	4.9	5.0	4.4	10.16

series and of y_0 will lead to different rejection frequencies. The upper part of the table gives sizes of asymptotic tests. $\mathscr{F}$ refers to the standard F test; $\mathscr{W}$ to the corresponding Wald test where no degrees of freedom corrections are employed and χ^2 critical values are used; $t<$ and $t>$ refer to one sided t tests against the alternatives $\lambda < \lambda_0$ and $\lambda > \lambda_0$ respectively. We find that the size distortions can be quite substantial. We did not attempt to obtain an upper bound to the actual sizes over the whole parameter space; possibly this upper bound is 1. From the limited set of simulations of the actual null distributions for $\lambda = \lambda_0$ of the statistic $\mathscr{F}$ over $(\lambda, \sigma_\varepsilon)$ we established a "size corrected" critical value (indicated by $\tilde{\mathscr{F}}^C$). Since we did not look at the whole parameter space (for each λ_0 only two values of σ_ε) and the null distribution of the standard F

statistic depends on both $(\lambda_0, \sigma_\varepsilon)$ and the exogenous regressor coefficients $(\beta_1, \beta_2, \beta_3)$, it appears most likely that these size corrections still involve over-rejections and thus power comparisons which may be heavily biased against the similar tests $\mathscr{L}^*$ and $\mathscr{L}^{**}$.

In the lower part of the table we present rejection frequencies for $\mathscr{F}^C$, which refers to the (incompletely) size-corrected F test, and for $\mathscr{L}^*$ and $\mathscr{L}^{**}$, the two exact (two-sided) tests. Estimated sizes for $\mathscr{L}^*$ and $\mathscr{L}^{**}$ may differ from 5% due to the randomness of the Monte Carlo. We find that neither of the three tests is superior. Test $\mathscr{L}^{**}$ seems to have higher power than $\mathscr{L}^*$ if $\lambda > \lambda_0$ (and vice versa). The exact tests prove to have power which may be as high as 100%, and their rejection frequency is occasionally found to be higher than for $\mathscr{F}^C$. Note that a genuinely size-corrected F test will have a rejection frequency not exceeding that of $\mathscr{F}^C$ and for use in practice it will be much more computer intensive than our procedures, since it requires a full search over the nuisance parameter space.

7. EMPIRICAL APPLICATIONS

For an application and illustration of the exact procedures in the context of pure univariate time-series analysis we used annual data published by Balke and Gordon (1986, pp. 781–786) on nominal *GNP* and *M2* in the US. We analyze the natural logarithm of *GNP/M2*, i.e. the velocity of *M2* (indicated by v_t), over the period 1959–1983. Estimation by OLS yields

$$\text{(7.1)} \qquad v_t = \underset{(0.037)}{0.135} - \underset{(0.020)}{0.005}\,(t/T) + \underset{(0.090)}{(0.723)}\,v_{t-1} + \hat{\varepsilon}_t \qquad (t = 1, \ldots, T),$$

$$T = 35; \quad s = 0.0286; \quad R^2 = 0.739; \quad DF = -3.078.$$

Estimated (asymptotic) standard errors are presented in parentheses; s is the usual estimator of σ and DF is the t ratio Dickey-Fuller statistic for testing $\lambda = 1$. Asymptotic tests for higher order (up to fourth) serial correlation, for structural change, and for non-normality of the disturbances have large p values and hence do not indicate severe misspecification. This is also the case for the more restricted model:

$$\text{(7.2)} \qquad v_t = \underset{(0.034)}{0.138} + \underset{(0.074)}{0.710}\,v_{t-1} + \hat{\varepsilon}_t \qquad (t = 1, \ldots, T);$$

$$T = 35; \quad s = 0.0281; \quad R^2 = 0.739; \quad DF = -3.919.$$

For producing exact inference we chose $\mathscr{D}_\lambda = [-1, +1]$ and used 999 replications for the Monte Carlo tests.

When assuming normality of the disturbances in specification (7.1) we found, upon using the $\mathscr{L}_\lambda^{**}$ statistic, an exact 95% confidence set for λ given by $\mathscr{C}_\lambda^{**}$ (0.05) = [0.64, 1.00]. The statistic $\mathscr{L}_\lambda^*$ yields the wider region $\mathscr{C}_\lambda^*$ (0.05) = [0.28, 1.00]. Hence we see that the corresponding unit root hypothesis tests (which are both in fact equivalent to particular Dickey-Fuller type tests) do not reject $\lambda = 1$. However, when we use our comprehensive procedures, and test in (7.4) the joint null hypothesis $\lambda = 1$ and $\beta_\tau = 0$ (we indicate the intercept coefficient by β_ι and the trend coefficient by β_τ), we find for both the statistics $\mathscr{L}^{**}_{(\lambda,\gamma)}(\lambda_0, \gamma_0)$ and $\mathscr{L}^{\dagger **}_{(\lambda,\gamma)}(\lambda_0, \gamma_0)$ p values of 0.042, and hence, the pure random walk with drift model is rejected at level 5%. Again the tests that do not include the zero restrictions on the redundant regressors under the null seem less powerful; they both yield a p value of 0.136 (note that conditioning on y_0 has apparently no effect here). If we do "a test on the regression," viz. $\lambda = 0$ and $\beta_\tau = 0$, all four tests mentioned in Proposition 1 yield zero p values. Testing the same joint hypotheses by two-stage procedures, the random walk with drift hypothesis is rejected at 10% level, since the p value of the conditional test statistic (4.1) is 0.043; zero restrictions on all coefficients apart from the intercept are rejected right away as $\lambda = 0$ is not in the confidence set $\mathscr{C}^{**}_\lambda(0.05)$.

Exploiting the $\mathscr{L}^{**}_\lambda$ interval we find the 90% confidence intervals [0.004, 0.190] and [-0.078, 0.040] for β_ι and β_τ respectively. Upon testing the significance of the trend we obtain a p value for $\mathscr{L}^L_\gamma(\gamma_0)$ of 0.994, so a zero value of the trend coefficient should certainly not be rejected (as we already learned from the confidence set), but we also find a p value for $\mathscr{L}^U_\gamma(\gamma_0)$ of 0.043; this is smaller than 0.15 and hence the bounds-test with level 10% ($\alpha_1 = 0.05 = \alpha_2$) is inconclusive. So, acceptance of a zero trend coefficient (which seems more or less self-evident following naive asymptotic reasoning) is in fact not strongly sustained by the 35 data observations.

Upon applying the exact procedures under normality to the parsimonious parameterization (7.2) we obtained $\mathscr{C}^{**}_\lambda(0.05) = [0.64, 0.99]$, so the unit root hypothesis is rejected now. This interval yields a 90% conservative confidence set [-0.001, 0.178] for the intercept, which thus is insignificant.

For illustration purpose (and not primarily for its empirical relevance) we also examined non-normal disturbances, viz. i.i.d. disturbances that are distributed as $(\chi^2_1 - 1)$, i.e. skew with zero mean, and t_1 (or Cauchy) disturbances, i.e., symmetric but with no finite moments. For the skew disturbances we found for (7.5) the interval $\mathscr{C}^{**}_\lambda(0.05) = [0.64, 0.96]$ and the intercept is now significantly positive, since $\mathscr{L}^U_\lambda(\lambda_0)$ has a p value of 0.004; thus the bounds test rejects. The Cauchy disturbances yield $\mathscr{C}^{**}_\lambda(0.05) = [0.61, 1.00]$ in (7.2) and the p values of $\mathscr{L}^U_\gamma(\gamma_0)$ and $\mathscr{L}^L_\gamma(\gamma_0)$ for the intercept are 0.380 and 0.001 respectively, so the bounds test does not reject; the conservative confidence set for the intercept is [-0.004, 0.191].

To illustrate the inference techniques in a simple structural econometric model, we built on an empirical study on narrow money demand in Canada by Marothia and Phillips (1982), henceforth MP. They perform various regressions in order to identify the key explanatory variables for money demand and supply functions, by using both single and simultaneous equation estimation techniques. In the case of the demand equation OLS, 2SLS, 3SLS, and IV give virtually the same results, viz. that real cash balance (m_t) is determined only by real income (y_t), the short-term interest rate (r_t), lagged real cash balance (m_{t-1}), and an intercept. MP's study is based on logs of quarterly data from

1970 (I) until 1979 (IV). During this period Canada had a flexible exchange rate regime. All data, except interest rates, are seasonally adjusted. The effects of long-term interest rates and wealth are found to be insignificant. The similarity of the various estimates is interpreted here as indicating that simultaneity is not a major issue for this particular equation. Of course, strong-exogeneity (especially of y_t) is highly unlikely here.

We adopt the final preferred specification of MP and obtain (due to data revisions our findings differ slightly from those published by MP):

$$m_t = \underset{(0.304)}{-\;0.430} + \underset{(0.039)}{0.083}\, y_t - \underset{(0.010)}{0.058}\, r_t + \underset{(0.056)}{0.913}\, m_{t-1} + \hat{\varepsilon}_t \quad (t = 1, \ldots, T); \tag{7.3}$$

$$T = 39; \quad s = 0.0136; \quad R^2 = 0.9594.$$

No obvious statistical evidence is found regarding structural breaks, (higher order) serial correlation, or non-normality of the disturbances. The main conclusions of MP on Canadian narrow money demand concern the estimated values of the long-run elasticities with respect to income and interest; equation (7.3) yields the plausible estimates 0.95 and -0.67 respectively.

We first examine standard asymptotic inference on the coefficient values as typically provided in current practice (β_1 is the intercept, β_2 and β_3 are the y and r coefficients respectively). Table II presents confidence intervals for individual coefficients obtained at nominal levels 95 and 90%. Such intervals can either be based on critical values of the standard Normal or the Student distribution. The latter (indicated by "asymptotic t") include a degrees of freedom correction, which—as is generally believed—should reduce the approximation errors committed in small samples. The (uncorrected) pure asymptotic intervals are indicated by "asymptotic N." The intervals for λ, although obtained by the standard OLS recipe, are not genuinely asymptotic intervals, since the normal asymptotic distribution is only valid for $|\lambda| < 1$; tests where $\lambda_0 \geq 1$ require different asymptotic null distributions.

Table III presents *exact* confidence intervals $\mathscr{C}_\lambda(\alpha_1)$ obtained under normality from $\mathscr{L}_\lambda^*$ and $\mathscr{L}_\lambda^{**}$ for $N = 2000$ (but $N = 1000$ gave virtually the same results). In the upper part of the table we present results for $\lambda \in \mathscr{D}_\lambda \equiv [-2, +2]$. Both exact intervals for λ are much wider than the asymptotic intervals.

TABLE II

ASYMPTOTIC CONFIDENCE INTERVALS FOR THE INDIVIDUAL COEFFICIENTS OF THE CANADIAN MONEY DEMAND EQUATION (7.3)

	Nominal level = 95%		Nominal level = 90%	
Coefficient	Asymptotic N	Asymptotic t	Asymptotic N	Asymptotic t
λ	[0.81, 1.02]	[0.80, 1.03]	[0.83, 1.00]	[0.82, 1.01]
β_1	[−0.99, 0.13]	[−1.05, 0.19]	[−0.90, 0.04]	[−0.94, 0.08]
β_2	[0.01, 0.16]	[0.00, 0.16]	[0.02, 0.14]	[0.02, 0.15]
β_3	[−0.08, −0.04]	[−0.08, −0.04]	[−0.07, −0.04]	[−0.07, −0.04]

TABLE III

EXACT CONFIDENCE INTERVALS FOR THE INDIVIDUAL COEFFICIENTS OF THE CANADIAN MONEY DEMAND EQUATION (7.3)

$\mathscr{D}_\lambda$	Coefficient	Confidence level	Interval	Statistic(s)
$[-2, +2]$	λ	95%	$[0.573, 1.363]$	$\mathscr{L}_\lambda^*$
		95%	$[0.716, 1.169]$	$\mathscr{L}_\lambda^{**}$
	β_1	$\geq 90\%$	$[-1.365, 0.645]$	$\mathscr{L}_\lambda^{**}$ and $\mathscr{L}_\gamma^L$
	β_2	$\geq 90\%$	$[-0.117, 0.246]$	
	β_3	$\geq 90\%$	$[-0.084, -0.029]$	
$[-1, +1]$	λ	95%	$[0.573, 1]$	$\mathscr{L}_\lambda^*$
		95%	$[0.716, 1]$	$\mathscr{L}_\lambda^{**}$
	β_1	$\geq 90\%$	$[-1.365, 0.292]$	$\mathscr{L}_\lambda^{**}$ and $\mathscr{L}_\gamma^L$
	β_2	$\geq 90\%$	$[-0.017, 0.246]$	
	β_3	$\geq 90\%$	$[-0.084, -0.036]$	

Apparently the information in the data or the capabilities of the tests are very weak with respect to rejecting exorbitant (from an economic point of view) λ values. The test $\mathscr{L}_\lambda^{**}$, which restricts the coefficients of the redundant variables under the null, seems the more effective one for these data. A fair comparison with the asymptotic intervals would require the assessment of the actual size of the latter type of intervals, which is impossible, due to the dependence on unknown nuisance parameter values.

The intervals for the β's have been obtained by employing (3.7) for $\alpha_1 = 0.05 = \alpha_2$ with $\mathscr{C}^{**}(0.05)$. The intervals for β indicate that, according to these exact procedures, the income coefficient is not significantly different from zero at the 10% level. As it seems, in this relationship and for these data, an estimated coefficient of λ greater than unity can easily be accommodated by a negative income effect. The results in the lower part of Table III, which are obtained under the a priori restriction $|\lambda| \leq 1$, still indicate insignificance of the income coefficient.

In the first row of Table IV joint significance of the explanatory variables is found by all four comprehensive tests and also by the two-stage procedure. In

TABLE IV

EXACT p-VALUES FOR VARIOUS JOINT HYPOTHESES ON THE COEFFICIENTS OF THE CANADIAN MONEY DEMAND EQUATION (7.3); $\mathscr{D}_\lambda = [-1, +1]$

	Test statistic(s)				
Null hypothesis	$\mathscr{L}_{\lambda,\gamma}^*$	$\mathscr{L}_{\lambda,\gamma}^{**}$	$\mathscr{L}_{\lambda,\gamma}^{\dagger}$	$\mathscr{L}_{\lambda,\gamma}^{\dagger *}$	$\mathscr{L}_\gamma^L$ & $\mathscr{C}_\lambda^{**}$
$\lambda = 0,\ \beta_2 = \beta_3 = 0$	0.000	0.000	0.000	0.000	0.000
$\lambda = 0.86,\ \beta_2 = 0.12,\ \beta_3 = -0.05$	0.253	0.392	0.253	0.392	0.268
$\lambda = 1,\ \beta_2 = \beta_3 = 0$	0.003	0.000	0.003	0.000	0.000
$\beta_2 = \beta_3 = 0$			not applicable		0.000
$\beta_2 + \beta_3 = 0$			not applicable		0.985

the second row the three values for the coefficients found by MP are tested and not rejected. Testing the (structural) MP specification against the pure time-series simple random walk with drift specification, i.e. $\lambda = 1$, $\beta_2 = \beta_3 = 0$, yields strong rejections at the 1% level. Hence, the explanatory variables suggested by economic theory produce a statistically highly significant improvement over the pure descriptive unit-root model.

Building on the confidence interval $\mathscr{C}_{\lambda}^{**}(0.05)$, procedure (3.11) yields a zero p-value for the hypothesis $\beta_2 = \beta_3 = 0$. Hence, it is rejected at level $\alpha = 0.05 + \alpha_2$ for any $\alpha_2 > 0$. For the hypothesis $\beta_2 + \beta_3 = 0$ this procedure yields an insignificant $\mathscr{L}_{\gamma}^{L}$ value. However, the test of this hypothesis is inconclusive since $\mathscr{L}_{\gamma}^{U}$ has a zero p value, which is below $\alpha_2' = 2\alpha_1 + \alpha_2 \geq 10\%$. Tests on long-run elasticities are sensible only when we restrict the domain of λ. If we choose $\mathscr{D}_{\lambda} = [0, 1)$, then the confidence set will be unbounded. Bounded sets can be obtained from (4.10) or (4.11) if we choose $\mathscr{D}_{\lambda} = [0, \lambda^U]$ with $\lambda^U < 1$, which is not pursued here.

8. CONCLUSION

By exploiting techniques to annihilate nuisance parameters from test statistics and by using these in combination with generalized bounds test procedures we are able to produce exact inference techniques on virtually any form of hypothesis on the coefficients of a first-order autoregressive distributed lag model. The resulting procedures are computer intensive, but nevertheless operational. As usual, genuine exactness is only obtained under specific conditions. In the present case the conditions are of similar nature as those for exactness of $\mathscr{t}$ and $\mathscr{F}$ tests in the static linear regression model. These conditions are: strongly exogenous regressors and i.i.d. normally distributed disturbances. In fact, the requirements for exactness of the techniques presented here for the first-order dynamic regression model are weaker, since we can deal straightaway with any (up to a scale factor known) form of distribution of the disturbances.

Our inference procedures can also be justified under weaker regularity conditions on the basis of asymptotic arguments; see Dufour and Kiviet (1994) for further details. Of course practitioners can take the inexactness of standard tests for granted and, although a finite sample is analyzed, rely simply on asymptotic theory assuming that particular regularity conditions are fulfilled and asymptotic properties do reasonably well hold for the actual finite sample under study. Here, however, we have shown how such speculations and sometimes quite inaccurate approximations (see also Nankervis and Savin (1985, 1987)) can be avoided. The nonsimilarity of the classical inference techniques in finite samples precludes the assessment of their actual significance and power in practical empirical situations, and a proper comparison of the efficiency of the two approaches is therefore hard to establish. This seems feasible to a limited degree in controlled simulation experiments only, as we demonstrated. When comparing a particular data distribution compatible with a given composite null hypothesis, such as $H_0 : \lambda = 1$, against some alternative distribution, an asymptotic test may appear to have more "power" because it rejects the null hypothesis more frequently under the alternative. But this is typically misleading: the asymptotic test should be "size-corrected" with respect to the *whole null hypothesis* to bring its *maximal* rejection probability under the null hypothesis at the nominal level. The appropriate size correction may require an important increase in the critical value. In the Monte Carlo experiments presented in Section 7, we could ascertain that the appropriate size corrections could be very substantial, but we could only perform partial ones. For other illustrations of this phenomenon, see Campbell and Dufour (1997) and Dufour and Torrès (1997). Currently there is no method for finding the theoretically correct critical value, and the methods proposed in this paper are the only ones which are provably valid for the inference problems considered in this article.

CRDE and Département de Sciences Économiques, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec H3C 3J7, Canada; dufour@plgcn.umontreal.ca

and

Tinbergen Institute and Faculty of Economics and Econometrics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands; jfk@fee.uva.nl.

Manuscript received November, 1994; final revision received March, 1997.

APPENDIX

PROOF OF PROPOSITION 1: Upon dividing the numerator and denominator of the various test statistics by σ^2 it is obvious that, under $\lambda = \lambda_0$ and $\gamma = \gamma_0$, the numerators are pivotal quadratic forms in ε/σ. Proceeding as in (2.9) the unrestricted sum of squared residuals of (2.15) and (2.17) can be expressed as

$$\varepsilon' M[W(\lambda_0,\gamma_0)]\varepsilon = \varepsilon' M[Z(\lambda_0,\gamma_0)]\varepsilon - \frac{\{\varepsilon' M[Z(\lambda_0,\gamma_0)]\mathbf{y}_{-1}\}^2}{\mathbf{y}_{-1}' M[Z(\lambda_0,\gamma_0)]\mathbf{y}_{-1}},$$

with $Z(\lambda_0,\gamma_0) = [Z \,\vdots\, \bar{Z} \,\vdots\, \iota(\lambda_0) \,\vdots\, C(\lambda_0)Z\gamma_0 \,\vdots\, C(\lambda_0)\bar{Z}]$. Under $\lambda = \lambda_0$, $\gamma = \gamma_0$ we find from (2.11) that $M[Z(\lambda_0,\gamma_0)]\mathbf{y}_{-1}/\sigma = M[Z(\lambda_0,\gamma_0)]C(\lambda_0)\varepsilon/\sigma$, and so both $\mathscr{L}^{*}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ and $\mathscr{L}^{**}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ yield pivotal test statistics. For test statistics (2.21) and (2.22), where $Z^{\dagger}(\lambda_0,\gamma_0) = [Z \,\vdots\, \bar{Z} \,\vdots\, y_0\iota(\lambda_0) + C(\lambda_0)Z\gamma_0 \,\vdots\, C(\lambda_0)\bar{Z}]$, we find $M[Z^{\dagger}(\lambda_0,\gamma_0)]\mathbf{y}_{-1}/\sigma = M[Z^{\dagger}(\lambda_0,\gamma_0)]C(\lambda_0)\varepsilon/\sigma$, where $Z^{\dagger}(\lambda_0,\gamma_0)$ depends on y_0, but if y_0 is known and fixed, both $\mathscr{L}^{\dagger}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ and $\mathscr{L}^{\dagger *}_{\lambda,\gamma}(\lambda_0,\gamma_0)$ are pivotal. *Q.E.D.*

PROOF OF PROPOSITION 2: Since $(\lambda,\gamma) \in \mathscr{C}_{\lambda,\gamma}(\alpha_1,\alpha_1)$ implies $\gamma \in \mathscr{C}^{U}_{\gamma}(\alpha_1,\alpha_2)$, we have $P[\gamma \in \mathscr{C}^{U}_{\gamma}(\alpha_1,\alpha_2)] \geq P[(\lambda,\gamma) \in \mathscr{C}_{\lambda,\gamma}(\alpha_1,\alpha_2)] \geq 1-\alpha$. Define now the set

$$\bar{\mathscr{C}}^{L}_{\gamma}(\alpha_1,\alpha_2) = \{\gamma_0 : \exists \lambda_0 \in \mathscr{C}_{\lambda}(\alpha_1) \text{ such that } (\lambda_0,\gamma_0) \notin \mathscr{C}_{\lambda,\gamma}(\alpha_1,\alpha_2)\}, \quad \text{if } \mathscr{C}_{\lambda}(\alpha_1) \neq \varnothing,$$
$$= \mathbb{R}^r, \quad \text{if } \mathscr{C}_{\lambda}(\alpha_1) = \varnothing.$$

This is the complement of $\mathscr{C}^{L}_{\gamma}(\alpha_1,\alpha_2)$ in the space of all admissible values of γ, hence $P[\gamma \in \mathscr{C}^{L}_{\gamma}(\alpha_1,\alpha_2)] = 1 - P[\gamma \in \bar{\mathscr{C}}^{L}_{\gamma}(\alpha_1,\alpha_2)]$. We also define

$$\mathscr{C}^{(1)}_{\lambda,\gamma}(\alpha_1,\alpha_2) = \{(\lambda_0,\gamma_0) : \lambda_0 \in \mathscr{C}_{\lambda}(\alpha_1) \text{ and } (\lambda_0,\gamma_0) \notin \mathscr{C}_{\lambda,\gamma}(\alpha_1,\alpha_2)\}.$$

Then

$$P[(\lambda,\gamma)\in\mathscr{C}^{(1)}_{\lambda,\gamma}(\alpha_1,\alpha_2)]=P[\lambda\in\mathscr{C}_\lambda(\alpha_1)\text{ and }\gamma\notin\mathscr{C}_{\gamma|\lambda}(\alpha_2)]$$
$$\geq 1-\alpha_1-(1-\alpha_2)=\alpha_2-\alpha_1,$$

where $\mathscr{C}_{\gamma|\lambda}(\alpha_2)$ is defined by (3.2). Note that $(\lambda,\gamma)\in\mathscr{C}^{(1)}_{\lambda,\gamma}(\alpha_1,\alpha_2)$ implies $\gamma\in\overline{\mathscr{C}}^{L}_{\gamma}(\alpha_1,\alpha_2)$, so that

$$P[\gamma\in\overline{\mathscr{C}}^{L}_{\gamma}(\alpha_1,\alpha_2)]\geq P[(\lambda,\gamma)\in\mathscr{C}^{(1)}_{\lambda,\gamma}(\alpha_1,\alpha_2)]\geq\alpha_2-\alpha_1.$$

Combining the above, we find $P[\gamma\in\mathscr{C}^{L}_{\gamma}(\alpha_1,\alpha_2)]\leq 1-\alpha_2+\alpha_1$. Consequently, a liberal confidence set for γ is found by replacing α_2 by $\alpha_2'=\alpha+\alpha_1$, giving $P[\gamma\in\mathscr{C}^{L}_{\gamma}(\alpha_1,\alpha_2')]\leq 1-\alpha$. Since $\alpha_2\leq\alpha\leq\alpha_2'$, we have $\mathscr{C}_{\lambda,\gamma}(\alpha_1,\alpha_2)\supseteq\mathscr{C}_{\lambda,\gamma}(\alpha_1,\alpha_2')$, hence $\mathscr{C}^{L}_{\gamma}(\alpha_1,\alpha_2')\subseteq\mathscr{C}^{L}_{\gamma}(\alpha_1,\alpha_2)\subseteq\mathscr{C}^{U}_{\gamma}(\alpha_1,\alpha_2)$. *Q.E.D.*

PROOF OF PROPOSITION 3: From the definitions (4.4) and (4.7), it follows that $S_1(\lambda)\geq S(\lambda)$. If $(\lambda,\beta'R')'\in\Gamma_0$ then the infimum $\bar{S}_0(\lambda,\gamma_0)$ is obtained over $(\lambda,\gamma_0')'\in\Gamma_0$ where $\gamma_0=\gamma=R\beta$, whereas $S_0(\lambda)$ is obtained over all $\gamma_0=R\beta_0\in\mathbb{R}^r$ obeying $(\lambda,\gamma_0')'\in\Gamma_0$, hence $\bar{S}_0(\lambda,\gamma_0)\geq S_0(\lambda)$. Thus, under the null hypothesis we have $S_0(\lambda)/S_1(\lambda)\leq\bar{S}_0(\lambda,\gamma_0)/S(\lambda)$, and so for any real x

$$P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)>x]=P\left[\frac{S_0(\lambda)}{S_1(\lambda)}>x\right]\leq P\left[\frac{\bar{S}_0(\lambda,\gamma_0)}{S(\lambda)}>x\right]=P[\mathscr{L}_{\gamma|\lambda}(\gamma_0)>x],$$

from which the conservative critical value easily follows. *Q.E.D.*

PROOF OF PROPOSITION 4: When $\lambda\in C_\lambda(\alpha_1)$ we have $\mathscr{L}^{L}_{\Gamma}(\Gamma_0,\Omega_1)\leq\mathscr{L}_{\Gamma|\gamma}(\Gamma_0,\Omega_1)$, and thus, for any real x,

$$P[\mathscr{L}^{L}_{\Gamma}(\Gamma_0,\Omega_1)>x]=P[\mathscr{L}^{L}_{\Gamma}(\Gamma_0,\Omega_1)>x\text{ and }\lambda\in C_\gamma(\alpha_1)]$$
$$+P[\mathscr{L}^{L}_{\Gamma}(\Gamma_0,\Omega_1)>x\text{ and }\lambda\notin C_\lambda(\alpha_1)]$$
$$\leq P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)>x\text{ and }\lambda\in C_\lambda(\alpha_1)]+P[\lambda\notin C_\lambda(\alpha_1)]$$
$$\leq P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)>x]+\alpha_1.$$

Hence, $P[\mathscr{L}^{L}_{\Gamma}(\Gamma_0,\Omega_1)>\tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2;X,R)]\leq\alpha_2+\alpha_1=\alpha$. Similarly, $\lambda\in C_\lambda(\alpha_1)$ implies $\mathscr{L}^{U}_{\Gamma}(\Gamma_0,\Omega_1)\geq\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)$, and

$$P[\mathscr{L}^{U}_{\Gamma}(\Gamma_0,\Omega_1)>x]\geq P[\mathscr{L}^{U}_{\Gamma}(\Gamma_0,\Omega_1)>x\text{ and }\lambda\in C_\lambda(\alpha_1)]$$
$$\geq P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)>x\text{ and }\lambda\in C_\lambda(\alpha_1)]$$
$$\geq 1-P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)\leq x]-P[\lambda\notin C_\lambda(\alpha_1)]$$
$$\geq 1-P[\mathscr{L}_{\Gamma|\lambda}(\Gamma_0,\Omega_1)\leq x]-\alpha_1.$$

Hence, $P[\mathscr{L}^{U}_{\Gamma}(\Gamma_0,\Omega_1)>\tilde{\mathscr{L}}_{\gamma|\lambda}(\alpha_2';X,R)]\geq\alpha_2'-\alpha_1=\alpha$. *Q.E.D.*

REFERENCES

BALKE, N. S., AND R. J. GORDON (1986): "Historical Data," in *The American Business Cycle*, ed. by R. J. Gordon. Chicago: The University of Chicago Press, 781–850.

BARNARD, G. A. (1963): "Discussion on M. S. Bartlett, The Spectral Analysis of Point Processes," *Journal of the Royal Statistical Society B*, 25, 294.

BIRNBAUM, Z. W. (1974): "Computers and Unconventional Test-Statistics," in *Reliability and Biometry*, ed. by F. Proschan and R. J. Serfling. Philadelphia: SIAM, 441–458.

CAMPELL, B., AND J.-M. DUFOUR (1997): "Exact Nonparametric Tests of Orthogonality and Random Walk in the Presence of a Drift Parameter," *International Economic Review*, 38, 151–173.

DICKEY, D. A., AND W. A. FULLER (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

DUFOUR, J.-M (1989): "Non-Linear Hypotheses, Inequality Restrictions, and Non-Nested Hypotheses: Exact Simultaneous Tests in Linear Regressions," *Econometrica*, 57, 335–355.

——— (1990): "Exact Tests and Confidence Sets in Linear Regressions with Autocorrelated Errors," *Econometrica*, 58, 479–494.

DUFOUR, J.-M., AND M. DAGENAIS (1985): "Durbin-Watson Tests for Serial Correlation in Regressions with Missing Observations," *Journal of Econometrics*, 27, 371–381.

DUFOUR, J.-M., AND M. HALLIN (1987): "Tests Non Paramétriques Optimaux pour le Modèle Autorégressif d'Ordre Un," *Annales d'Economie et de Statistique*, 5, 411–434.

DUFOUR, J.-M., AND M. L. KING (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., AND J. F. KIVIET (1994): "Exact Inference Methods for First-Order Autoregressive Distributed Lag Models," Tinbergen Institute Discussion Paper TI 94-143, and Cahier 4295, C.R.D.E., Université de Montréal.

DUFOUR, J.-M., AND O. TORRÈS (1997): "Union-Intersection and Sample-Split Methods in Econometrics with Applications to SURE and MA Models," in *Handbook of Applied Economic Statistics*, ed. by D. E. A. Giles and A. Ullah. New York: Marcel Dekker, forthcoming.

DWASS, M. (1957): "Modified Randomization Tests for Nonparametric Hypotheses," *Annals of Mathematical Statistics*, 28, 181–187.

EDGINGTON, E. S. (1980): *Randomization Tests.* New York: Marcel Dekker.

EFRON, B, (1979): "Bootstrap Methods: Another Look at the Jacknife," *The Annals of Statistics*, 7, 1–26.

EFRON, B., AND R. J. TIBSHIRANI (1993): *An Introduction to the Bootstrap.* New York: Chapman & Hall.

FOUTZ, R. V. (1980): "A Method for Constructing Exact Tests from Test Statistics that have Unknown Null Distributions," *Journal of Statistical Computation and Simulation*, 10, 187–193.

GIERSBERGEN, N. P. A. VAN, AND J. F. KIVIET (1996): "Bootstrapping a Stable AD Model: Weak vs. Strong Exogeneity," *Oxford Bulletin of Economics and Statistics*, 58, 631–656.

GOURIÉROUX, C., AND A. MONFORT (1989): *Statistique et Modèles Économétriques.* Paris: Economica.

HÁJEK, J., AND Z. ŜIDÁK (1967): *Theory of Rank Tests.* New York: Academic Press.

HALL, P. (1992): *The Bootstrap and Edgeworth Expansions.* New York: Springer-Verlag.

——— (1994): "Methodology and Theory of the Bootstrap," Chapter 39 in *Handbook of Econometrics*, Vol. IV, ed. by R. F. Engle and D. L. McFadden. Amsterdam: Elsevier.

HALL, P., AND D. M. TITTERINGTON (1989): "The Effect of Simulation Order on Level Accuracy and Power of Monte Carlo Tests," *Journal of the Royal Statistical Society, Series B*, 51, 459–467.

HENDRY, D. F., AND G. MIZON (1978): "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Bank of England Demand-for-Money Function," *Economic Journal*, 88, 549–563.

HILLIER, G. H. (1987): "Classes of Similar Regions and Their Power Properties for Some Econometric Testing Problems," *Econometric Theory*, 3, 1–44.

HILLIER, G. H., AND M. L. KING (1987): "Linear Regression with Correlated Errors: Bounds on Coefficient Estimates and *t*-Values" in *Specification Analysis in the Linear Model: Essays in Honour of Donald Cochrane*, ed. by M. L. King and D. E. A. Giles. London: Routledge and Kegan Paul, 74–80.

JÖCKEL, H. K. (1986): "Finite Sample Properties and Asymptotic Efficiency of Monte Carlo Tests," *Annals of Statistics*, 14, 336–347.

KIVIET, J. F. (1980): "Effects of ARMA Errors on Tests for Regression Coefficients: Comments on Vinod's Article; Improved and Additional Results," *Journal of the American Statistical Association*, 75, 353–358.

——— (1991): "Tighter Bounds for the Effects of ARMA Disturbances on Tests for Regression Coefficients" in *Econometric Decision Models: New Methods of Modeling and Applications*, ed. by J. Gruber. Berlin: Springer-Verlag, 404–418.

KIVIET, J. F., AND G. D. A. PHILLIPS (1990): "Exact Similar Tests for the Root of a First-Order Autoregressive Regression Model," University of Amsterdam, AE Report 12/90, paper presented at Econometric Society World Meeting 1990, Barcelona, Spain.

——— (1992): "Exact Similar Tests for Unit Roots and Cointegration," *Oxford Bulletin of Economics and Statistics*, 54, 349–367.

KRÄMER, W., J. F. KIVIET, AND J. BREITUNG (1991): "True Versus Nominal Size of the F-Test in the Linear Regression Model with Autocorrelated Disturbances" in *Econometric Decision Models: New Methods of Modeling and Applications*, ed. by J. Gruber. Berlin: Springer-Verlag, 419–428.

MAROTHIA, D. K., AND W. E. PHILLIPS (1982): "Demand and Supply Functions for Money in Canada; Some Further Evidence," *Journal of Monetary Economics*, 9, 249–261.

NANKERVIS, J. C., AND N. E. SAVIN (1985): "Testing the Autoregressive Parameter with the t Statistic," *Journal of Econometrics*, 27, 143–161.

——— (1987): "Finite Sample Distributions of t and F Statistics in an AR(1) Model with an Exogenous Variable," *Econometric Theory*, 3, 387–408.

SARGAN, J. D. (1980): "Some Tests of Dynamic Specification for a Single Equation," *Econometrica*, 48, 879–897.

SCHMIDT, P. (1976): *Econometrics.* New York: Marcel Dekker.

THEIL, H., AND J. S. SHONKWEILER (1986): "Monte Carlo Tests of Autocorrelation," *Economics Letters*, 20, 157–160.

VINOD, H. D. (1976): "Effects of ARMA Errors on the Significance Tests for Regression Coefficients," *Journal of the American Statistical Association*, 71, 929–933.

VINOD, H. D., AND A. ULLAH (1981): *Recent Advances in Regression Methods.* New York: Marcel Dekker.

ZINDE-WALSH, V., AND A. ULLAH (1987): "On Robustness of Tests of Linear Restrictions in Regression Models with Elliptical Error Distributions," in *Time Series and Econometric Modelling, Advances in the Statistical Sciences*, Vol. III, ed. by I. B. MacNeill and G. J. Umphrey. Dordrecht: D. Reidel, 235–251.