OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 65, SUPPLEMENT (2003) 0305-9049

Exact Skewness–Kurtosis Tests for Multivariate Normality and Goodness-of-Fit in Multivariate Regressions with Application to Asset Pricing Models*

JEAN-MARIE DUFOUR†, LYNDA KHALAF‡ and MARIE-CLAUDE BEAULIEU§

†Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Centre interuniversitaire de recherche en économie quantitative (CIREQ), and Département de sciences économiques, Université de Montréal, Montréal, Québec, Canada H3C 3J7 (e-mail: jean.marie.dufour@umontreal.ca)
‡Centre interuniversitaire de recherche en économie quantitative (CIREQ), and GREEN, Université Laval, Québec, Canada, G1K 7P4 (e-mail: lynda.khalaf@ecn.ulaval.ca)
§Centre interuniversitaire de recherche sur les politiques économiques et l'emploi (CIRPÉE), Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Département de finance et assurance, Université Laval, Québec, Canada G1K 7P4 (e-mail: marie-claude.beaulieu@fas.ulaval.ca)

Abstract

We study the problem of testing the error distribution in a multivariate linear regression (MLR) model. The tests are functions of appropriately standardized multivariate least squares residuals whose distribution is invariant to the

*The authors thank two anonymous referees and the Editor Niels Haldrup for several useful comments. This work was supported by the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Alexander-von-Humboldt Foundation (Germany), the Institut de finance mathématique de Montréal (IFM2), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Québec), and the Fonds de recherche sur la nature et les technologies (Québec). This paper was also partly written at the Centre de recherche en économie et statistique (INSEE, Paris), the Technische Universität Dresden (Fakultät Wirtschaftswissenschaften) and the Finance Division at the University of British Columbia.

JEL Classification numbers: C3, C12, C33, C15, G1, G12, G14.

unknown cross-equation error covariance matrix. Empirical multivariate skewness and kurtosis criteria are then compared with a simulation-based estimate of their expected value under the hypothesized distribution. Special cases considered include testing multivariate normal and stable error distributions. In the Gaussian case, finite-sample versions of the standard multivariate skewness and kurtosis tests are derived. To do this, we exploit simple, double and multi-stage Monte Carlo test methods. For non-Gaussian distribution families involving nuisance parameters, confidence sets are derived for the nuisance parameters and the error distribution. The tests are applied to an asset pricing model with observable risk-free rates, using monthly returns on New York Stock Exchange (NYSE) portfolios over 5-year subperiods from 1926 to 1995.

I. Introduction

Drawing inference from the parameters of multivariate linear regression (MLR) models is a basic statistical problem. Such models, which can combine both cross-section and time series data, are common in various fields of statistics and econometrics (see Stewart, 1997; Dufour and Khalaf, 2002a, b; 2003, and the references therein). Important cases include consumer and factor demand systems, reduced forms derived from linear simultaneous equation models, and various asset pricing models in finance.[1]

Tests and confidence sets for MLR models are heavily influenced by the disturbance distribution and the assumptions made on the latter. Although standard asymptotic procedures are distribution-free, their performance in finite samples tends to be quite bad (see, for example, Dufour and Khalaf, 2002a, b). Some finite-sample results are available in the statistical literature, but these are almost entirely restricted to Gaussian models. Another avenue which allows to obtain exact tests uses simulation-based tests, as described in Dufour and Khalaf (2002b). This approach allows one to relax normality, but still requires the formulation of a parametric model on the errors.[2] This situation underscores the importance of testing disturbance normality as well as other parametric distributional assumptions. Furthermore, specific distributional assumptions may be required by important economic or financial hypotheses, e.g. mean–variance efficiency in the context of the CAPM model (see Beaulieu, Dufour and Khalaf, 2001).

[1]Familiar MLR-based applications in finance include *market models*, such as the capital asset pricing model (CAPM) which may be traced back to Gibbons (1982). The associated empirical literature is enormous (for reviews, see Shanken, 1996; Campbell, Lo and MacKinlay, 1997).

[2]Heavy-tailed distributions important in financial modelling may easily be accommodated in this way.

In multivariate contexts, relatively little work has been done on testing distributional goodness-of-fit (GF) compared with the univariate case. This holds even when the hypothesized null distribution is multivariate normal (see the reviews of Mardia, 1980; D'Agostino and Stephens, 1986; Thode, 2002). Indeed, system diagnostic tests raise problems not encountered in the analysis of univariate models. A major difficulty arises from cross-equation disturbance correlations. Although it is desirable to use test procedures that take account of these correlations, they still constitute (unknown) nuisance parameters. The typical approach to this problem is to consider statistics whose distribution is asymptotically free of nuisance parameters (see, for example, Mardia, 1970; Richardson and Smith, 1993; Kilian and Demiroglu, 2000; Fiorentini, Sentana and Calzolari, 2003; and the review of Thode, 2002). However, in systems with many equations where the number of nuisance parameters is quite large relative to the sample size, such approximations are highly unreliable (see Horswell and Looney, 1992, 1993; Holgersson and Shukur, 2001).

The statistical literature on GF tests (see Mardia, 1980; Lütkepohl and Theilen, 1991; Horswell and Looney, 1992, 1993; Henze, 1994) has focused mainly on the location-scale model, which may be seen as a special case of the MLR model where the regressors reduce to a vector of ones. Indeed, the presence of covariates considerably complicates the testing problem and related (exact and asymptotic) distributional theory, even in univariate regressions (see Dufour *et al*., 1998 and the references therein). Despite the widespread recognition of such problems, our review of the statistics and econometrics literature has revealed that exact multivariate GF tests are unavailable, even for the Gaussian hypothesis or the location-scale model.

In this paper, we propose a general exact method for GF testing in MLR models. Our results can be summarized as follows. First, we address the distributional complications arising from the presence of covariates and unknown error covariances. We show that tests which use properly standardized residuals have a null distribution that does not depend on either regression coefficients, error variances or covariances, once the error distribution is parametrically specified up to an (unknown) linear transformation (or covariance matrix). More specifically, these tests are based on exploiting invariance properties for two distinct families of empirically scaled residuals: (1) a properly re-scaled version of the residual matrix (using the Cholesky root of the empirical residual covariance matrix); (2) the projector matrix associated with the least square residual matrix from the MLR model.[3]

[3]Although related pivotality results have been pointed out for the simpler Gaussian location-scale models (see Mardia, 1980; Lütkepohl and Theilen, 1991; Horswell and Looney, 1993), it does not appear those presented here have been used in the earlier literature on inference in general MLR models.

Secondly, we exploit the above invariance results to derive finite-sample tests of multinormality for the disturbances of MLR models. We consider two categories of test statistics based on empirical multivariate skewness and kurtosis coefficients: (1) multivariate extensions of the familiar Jarque–Bera tests [Jarque and Bera (1987, henceforth JB)], obtained by combining individual residual-based JB tests computed from individual equations (as suggested by Kilian and Demiroglu, 2000); (2) Mardia-type statistics (Mardia, 1970) based on empirical skewness and kurtosis derived from the least squares residual projector. These statistics have finite-sample null distributions which may be very difficult to evaluate through analytical methods. However, because of the fact that their distributions are free of nuisance parameters and easy to simulate, we can exploit the technique of Monte Carlo (MC) tests (Dwass, 1957; Barnard, 1963; Dufour and Kiviet, 1998; Dufour and Khalaf, 2001). This simulation-based procedure yields an exact test when the distribution of the test statistic is *pivotal* under the null hypothesis. As a result of the flexibility of this method, we define new methods for combining excess skewness and kurtosis criteria.

Thirdly, we show that the proposed multinormality tests can be easily adapted to assess other hypothesized disturbance distributions. For that purpose, the statistics are modified in order to compare empirical multivariate skewness and kurtosis measures with *simulation-based estimates* of their expected values under the hypothesized distribution (instead of theoretical, possibly inaccurate, expected values that may be difficult to derive). These corrections are also applicable in the Gaussian case. The MC test method then works in this case as in the previous one to achieve perfect size control, taking account of the fact that the simulated test statistics are exchangeable (because of the presence of simulated moment estimates) rather than independent identically distributed (i.i.d.), leading to a *double MC test* procedure. As long as the disturbance distribution is specified up to an unknown linear transformation, there is no restriction on the form of the tested distribution.[4] A *triple MC test* method is also proposed to combine several tests into an omnibus GF test.

Fourthly, the tests proposed are applied to an asset pricing model with observable risk-free rates. We assess multivariate normality and symmetric stable (possibly heavy tailed) distributions. The latter non-normal family involves a nuisance parameter (which defines tail-thickness) that we take as unknown. The proposed MC non-Gaussian GF tests are exact when the null hypothesis sets this parameter to specific values. On assembling the nuisance parameter values which are not rejected (i.e. by 'inverting' the GF tests), this yields *confidence sets* for the fitting distributions. Such confidence sets may

[4]The latter can be heavy-tailed and may even miss moments. The fact that a distribution does have a finite fourth moment does not preclude one to use an *empirical kurtosis* for assessing its GF.

then be used as an intermediate step in the context of other inference problems.

The paper is organized as follows. Section II describes the framework. In section III, we define standardized residuals, discuss their invariance properties, and state our basic exact distributional results. In section IV, we propose our multivariate GF test procedures. Section V reports our empirical analysis. We conclude in section VI.

II. Framework

Let us consider a system of correlated regression equations of the form:

$$Y = XB + U \tag{1}$$

where $Y = [Y_1, \ldots, Y_n]$ is a $T \times n$ matrix of observations on n dependent variables, X is a $T \times k$ full-column rank matrix, $B = [B_1, \ldots, B_n]$ is a $k \times n$ matrix of unknown fixed coefficients and $U = [U_1, \ldots, U_n] = [V_1, \ldots, V_T]'$ is a $T \times n$ matrix of random disturbances. Following Dufour and Khalaf (2002b), we suppose the errors have the following structure:

$$\begin{aligned} V_t = JW_t, \quad t = 1, \ldots, T, \\ w \equiv \text{vec}(W_1, \ldots, W_T) \sim \mathscr{F}(v), \end{aligned} \tag{2}$$

where $\mathscr{F}(v)$ is a known distribution, which may depend on the parameter v, and J satisfies one of the two following conditions:

$$J \text{ is an unknown non-singular lower triangular matrix;} \tag{3a}$$

$$J \text{ is an unknown non-singular matrix.} \tag{3b}$$

On setting $W = [W_1, \ldots, W_T]'$, the above assumptions entail that

$$W = U(J^{-1})'. \tag{4}$$

In particular, this condition will be satisfied for i.i.d. normal errors; in this case, assumptions (3a) and (3b) are equivalent. Let $\Sigma = JJ'$ which gives the covariance of V_t when $\text{Cov}(W_t) = I_n$. The least squares estimate of B and the corresponding residuals are

$$\hat{B} = (X'X)^{-1}X'Y, \qquad \hat{U} = \left[\hat{U}_1, \ldots, \hat{U}_n\right] = MU, \quad M = I - X(X'X)^{-1}X'. \tag{5}$$

It is clear from equations (4) and (5) that the distribution of $\hat{U}$ in general depends on the unknown scaling matrix J (or on the covariance matrix $\Sigma = JJ'$) so that test statistics based on $\hat{U}$ may involve J as a nuisance parameter. This will be the case in particular for the off-diagonal parameters which typically determine the cross-correlations of disturbances.

III. Multivariate standardized residuals

We now consider the problem of building residual-based test statistics whose null distribution will not be affected by the unknown scaling matrix J. In order to do this, we shall now state two general invariance results ensuring that appropriately standardized residuals have distributions which do not depend on J. Let

$$\tilde{W} = \hat{U} S_{\hat{U}}^{-1} \tag{6}$$

where $S_{\hat{U}}$ is the Cholesky factor of $\hat{U}'\hat{U}$, i.e. $S_{\hat{U}}$ is the (unique) upper triangular matrix such that $\hat{U}'\hat{U} = S_{\hat{U}}' S_{\hat{U}}$, $\left(\hat{U}'\hat{U}\right)^{-1} = S_{\hat{U}}^{-1}\left(S_{\hat{U}}^{-1}\right)'$.

Theorem 1. INVARIANCE OF CHOLESKY STANDARDIZED MULTIVARIATE RESIDUALS. Under equation (1) and for all error distributions compatible with equations (2) and (3a), the standardized residual matrix $\tilde{W}$ defined in equation (6) satisfies the identity

$$\tilde{W} = \hat{U} S_{\hat{U}}^{-1} = \hat{W} S_{\hat{W}}^{-1} \tag{7}$$

where $\hat{W} = MW$ and $S_{\hat{W}}$ is the (unique) upper triangular matrix such that

$$\hat{W}'\hat{W} = S_{\hat{W}}' S_{\hat{W}}, \quad \left(\hat{W}'\hat{W}\right)^{-1} = S_{\hat{W}}^{-1}\left(S_{\hat{W}}^{-1}\right)'. \tag{8}$$

Proof. Using equations (4), (5) and (8), we have:

$$\tilde{W} = \hat{U}(J^{-1})'(J'S_{\hat{U}}^{-1}) = MU(J^{-1})'(J'S_{\hat{U}}^{-1}) = MW\left(J'S_{\hat{U}}^{-1}\right) \tag{9}$$

$$\left(J'S_{\hat{U}}^{-1}\right)\left(J'S_{\hat{U}}^{-1}\right)' = J'S_{\hat{U}}^{-1}\left(S_{\hat{U}}^{-1}\right)'J = \left[(J^{-1})\hat{U}'\hat{U}(J^{-1})'\right]^{-1} \tag{10}$$

$$= \left[(J^{-1})U'MU(J^{-1})'\right]^{-1} = (W'MW)^{-1} = S_{\hat{W}}^{-1}\left(S_{\hat{W}}^{-1}\right)'. \tag{11}$$

On observing that $J'S_{\hat{U}}^{-1}$ is lower triangular, this means that $\left(J'S_{\hat{U}}^{-1}\right)'$ is the (unique) Cholesky factor of $(W'MW)^{-1}$, hence $J'S_{\hat{U}}^{-1} = S_{\hat{W}}^{-1}$. Substituting the latter identity in equation (9), we see that $\tilde{W} = MW\left(J'S_{\hat{U}}^{-1}\right) = \hat{W} S_{\hat{W}}^{-1}$. ■

Consider now the Mahalanobis matrix (see Mardia, 1970):

$$\hat{D} = \hat{U}\left(\hat{U}'\hat{U}/T\right)^{-1}\hat{U}'. \tag{12}$$

Theorem 2. INVARIANCE OF MAHALANOBIS RESIDUAL MATRIX. Under equation (1) and for all error distributions compatible with equations (2) and (3b), the residual-based Mahalanobis matrix $\hat{D}$ defined in equation (12) satisfies

$$\hat{D} = T\hat{W}\left(\hat{W}'\hat{W}\right)^{-1}\hat{W}'. \tag{13}$$

Proof. Using the identities $\hat{U} = MU$ and $U = WJ'$, we see that:

$$\begin{aligned}\hat{U}(\hat{U}'\hat{U}/T)^{-1}\hat{U}' &= TMU(U'MU)^{-1}U'M = TMU(J^{-1})'J'(U'MU)^{-1}JJ^{-1}U'M \\ &= TMU(J^{-1})'\left[(J^{-1})U'MU(J^{-1})'\right]^{-1}J^{-1}U'M \\ &= TMW(W'MW)^{-1}W'M = T\hat{W}\left(\hat{W}'\hat{W}\right)^{-1}\hat{W}'. \qquad \blacksquare\end{aligned}$$

The latter result relates to Theorem 1 as it is easy to see that $\hat{D} = T\hat{U}(\hat{U}'\hat{U})^{-1}\hat{U}' = T\hat{U}S_{\hat{U}}^{-1}(S_{\hat{U}}^{-1})'\hat{U}' = T\tilde{W}\tilde{W}'$. Theorems 1 and 2 include, as special cases, several known exact invariance results in the Gaussian location-scale model (see, for example, Mardia, 1970; Lütkepohl and Theilen, 1991; Thode, 2002).

IV. Skewness–kurtosis goodness-of-fit tests

We now use the above results to derive GF tests based on multivariate skewness and kurtosis coefficients. The proposed tests are formally valid for any parametric null hypothesis which takes the general form (2). Let us first consider the null hypothesis (2) where $\nu = \nu_0$ with ν_0 known.

Basic test statistics

The GF test statistics suggested here use two popular multivariate skewness and kurtosis measures. Specifically, we first consider extensions of the statistics

$$SK_{\mathrm{M}} = \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\hat{d}_{st}^3, \quad KU_{\mathrm{M}} = \frac{1}{T}\sum_{t=1}^{T}\hat{d}_{tt}^2, \tag{14}$$

where the variables $\hat{d}_{st}$ are the elements of the matrix $\hat{D} = \left[\hat{d}_{st}\right]$. These criteria were introduced by Mardia (1970) to assess deviations from multivariate normality, in location-scale models; see also Zhou (1993). Mardia further proposed the omnibus normality test:

$$MSK = \frac{T}{6}SK_{\mathrm{M}} + \frac{T[KU_{\mathrm{M}} - n(n+2)]^2}{8n(n+2)} \underset{T\to\infty}{\sim} \chi^2\left((n/6)(n+1)(n+2)+1\right) \tag{15}$$

where the symbol $\underset{T\to\infty}{\sim}$ refers to the asymptotic null distribution of the test statistic. Secondly, we consider extensions of the aggregate criteria applied by Kilian and Demiroglu (2000):

$$SK_{\mathrm{KD}} = (sk_1, \ldots, sk_n)'(sk_1, \ldots, sk_n), \tag{16}$$

$$KU_{\mathrm{KD}} = (ku_1 - 3, \ldots, ku_n - 3)'(ku_1 - 3, \ldots, ku_n - 3), \tag{17}$$

$$sk_i = \frac{T^{-1}\sum_{t=1}^{T}\tilde{W}_{ti}^3}{\left(T^{-1}\sum_{t=1}^{T}\tilde{W}_{ti}^2\right)^{3/2}}, \qquad ku_i = \frac{T^{-1}\sum_{t=1}^{T}\tilde{W}_{ti}^4}{\left(T^{-1}\sum_{t=1}^{T}\tilde{W}_{ti}^2\right)^{2}}, \quad i = 1, \ldots, n, \quad (18)$$

where $\tilde{W}_{ti}$ denotes the elements of the matrix $\tilde{W}$ defined by equation (6). In other words, sk_i and ku_i are the individual skewness and kurtosis measures based on the standardized residuals matrix (see Jarque and Bera, 1987; Kiefer and Salmon, 1983 for tests based on higher moments). The omnibus normality test studied by Kilian and Demiroglu (2000) is:

$$JB = \frac{T}{6}SK_{\mathrm{KD}} + \frac{T}{24}KU_{\mathrm{KD}} \underset{T\to\infty}{\sim} \chi^2(2n). \quad (19)$$

Extension to testing non-Gaussian distributions

To extend the above criteria beyond the Gaussian context, we shall modify them in three ways. First, we use these measures in excess of their expected values under equation (2). Secondly, we show that for given v, our modified test statistics are pivotal under the null hypothesis which allows to derive an exact simulation based p-value. Finally, we propose an exact combined skewness–kurtosis test. Our approach rests on the following invariance properties.

Theorem 3. Distribution of JB-type statistics in MLR. Under equation (1) and for all error distributions compatible with equations (2) and (3a), the multivariate skewness and kurtosis criteria sk_i and ku_i, $i = 1, \ldots, n$, defined in equation (18) are distributed like

$$\widetilde{sk}_i = \frac{T^{-1}\sum_{t=1}^{T}\bar{W}_{ti}^3}{\left(T^{-1}\sum_{t=1}^{T}\bar{W}_{ti}^2\right)^{3/2}}, \qquad \widetilde{ku}_i = \frac{T^{-1}\sum_{t=1}^{T}\bar{W}_{ti}^4}{\left(T^{-1}\sum_{t=1}^{T}\bar{W}_{ti}^2\right)^{2}}, \quad i = 1, \ldots, n, \quad (20)$$

respectively, where the $\bar{W}_{ti}$ are the elements of the matrix $\hat{W}S_{\hat{W}}^{-1}$ where $\hat{W} = MW$ and $S_{\hat{W}}$ is the Cholesky factor of $\hat{W}'\hat{W}$ as defined in equation (8), $M = I - X(X'X)^{-1}X'$, and W is defined by equation (2).

Theorem 4. Distribution of Mardia-type skewness and kurtosis. Under (1) and for all error distributions compatible with equations (2) and (3b), the multivariate skewness and kurtosis criteria SK_{M} and KU_{M} defined in equation (14) are distributed, like $\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}d_{st}^3$ and $\frac{1}{T}\sum_{t=1}^{T}d_{tt}^2$, respectively, where d_{st} are the elements of the matrix $T\hat{W}\left(\hat{W}'\hat{W}\right)^{-1}\hat{W}'$, $\hat{W} = MW$, $M = I - X(X'X)^{-1}X'$, and W is defined by equation (2).

The proof of both theorems follows immediately from Theorems 1 and 2. On this basis, we propose the following skewness–kurtosis-based statistics to test equation (2). Let

$$ESK_{\mathrm{M}}(v_0) = \left|SK_{\mathrm{M}} - \overline{SK}_{\mathrm{M}}(v_0)\right|, EKU_{\mathrm{M}}(v_0) = \left|KU_{\mathrm{M}} - \overline{KU}_{\mathrm{M}}(v_0)\right|, \quad (21)$$

$$ESK_{\mathrm{KD}}(v_0) = \left(esk_1(v_0), \ldots, esk_n(v_0)\right)'\left(esk_1(v_0), \ldots, esk_n(v_0)\right), \quad (22)$$

$$EKU_{\mathrm{KD}}(v_0) = \left(eku_1(v_0), \ldots, eku_n(v_0)\right)'\left(eku_1(v_0), \ldots, eku_n(v_0)\right), \quad (23)$$

$$esk_i(v_0) = \left(sk_i(v_0) - \mu_{sk_i(v_0)}\right)/\sigma_{sk_i(v_0)}, \ i = 1, \ldots, n, \quad (24)$$

$$eku_i(v_0) = \left(ku_i(v_0) - \mu_{ku_i(v_0)}\right)/\sigma_{ku_i(v_0)}, \ i = 1, \ldots, n, \quad (25)$$

where $\overline{SK}_{\mathrm{M}}(v_0)$ and $\overline{KU}_{\mathrm{M}}(v_0)$ are simulation-based estimates of the mean of SK_{M} and KU_{M} given equation (2), $\mu_{sk_i}(v_0)$ and $\sigma_{sk_i}(v_0)$ are simulation-based estimates of the mean and standard deviation of $sk_i(v_0)$ under equation (2), $ku_i(v_0)$ and $\sigma_{ku_i}(v_0)$ are simulation-based estimates of the mean and the standard deviation of $ku_i(v_0)$ given equation (2). For ease of presentation, we shall call these estimates '*reference simulated moments*' (RSM) and define

$$E = [ESK_{\mathrm{M}}(v_0), EKU_{\mathrm{M}}(v_0), ESK_{\mathrm{KD}}(v_0), EKU_{\mathrm{KD}}(v_0)]'. \quad (26)$$

To obtain these RSM, one may proceed as follows:

A1. Draw N_0 realizations of W following the distribution $\mathscr{F}(v_0)$ in equation (2);
A2. For each draw, construct the pivotal quantities $\hat{W}S_{\hat{W}}^{-1}$ and $T\hat{W}\left(\hat{W}'\hat{W}\right)^{-1}\hat{W}'$ which yield, applying Theorems 3 and 4, N_0 realizations of the statistics under consideration;
A3. The empirical moments of the latter simulated series yield the desired estimates.

Non-standard null distributions and multi-stage MC tests

Our modified test criteria have nonstandard null distributions, even under normal null hypotheses. Yet these distributions are pivotal (in normal and non-normal contexts) and can be easily simulated which justifies the application of Monte Carlo tests (Dufour, 2002). This procedure yields a bootstrap-type exact test when the null distribution of the underlying statistic is nuisance parameter free. When the statistic considered, say $S = S(Y, X)$ can be written as $S(Y, X) = \bar{S}(W, X)$, where Y and X are as in equation (1), W is defined by equation (2) and is fully specified, the general MC test methodology proceeds as follows.

Let $S^{(0)}$ denote the test statistic calculated from the observed data set; generate N replications $S^{(1)}, \ldots, S^{(N)}$ of the test statistic S so that $S^{(0)}, S^{(1)}, \ldots, S^{(N)}$ are exchangeable. Given the latter series, compute $\hat{p}_N[S] \equiv p_N(S^{(0)}; S)$, where

$$p_N(x;S) \equiv \frac{NG_N(x;S)+1}{N+1}, \quad G_N(x;S) \equiv \frac{1}{N}\sum_{i=1}^{N} s\big(S^{(i)} - x\big), \tag{27}$$

and $s(x) = 1$ if $x \geq 0$, and $s(x) = 0$ if $x < 0$. In other words, $p_N(S^{(0)}; S) = [N\hat{G}_N(S^{(0)}) + 1]/(N + 1)$ where $N\hat{G}_N(S^{(0)})$ is the number of simulated values greater than or equal to $S^{(0)}$. The MC critical region is: $p_N(S^{(0)}; S) \leq \alpha$, $0 < \alpha < 1$. If the distribution of S is continuous and $\alpha(N + 1)$ is an integer, then $P\big[p_N(S^{(0)}; S) \leq \alpha\big] = \alpha$ under the null hypothesis. When applied to the above GF criteria, the MC test technique can be summarized as follows.

B1. We obtain the RSM (according to A1–A3), which are generated only once, so the next steps are conditional on these estimates.

B2. Using the RSM and applying (21)–(23) to the data, we find the observed value of E:

$$E^{(0)} = [ESK_{\mathrm{M}}^{(0)}(v_0), EKU_{\mathrm{M}}^{(0)}(v_0), ESK_{\mathrm{KD}}^{(0)}(v_0), EKU_{\mathrm{KD}}^{(0)}(v_0)]'. \tag{28}$$

B3. Independently of the RSM and $E^{(0)}$, we draw N i.i.d. realizations of W according to $\mathscr{F}(v_0)$ in equation (2), and for each of these draws, we compute the pivotal quantities $\hat{W}S_{\hat{W}}^{-1}$ and $T\hat{W}\big(\hat{W}'\hat{W}\big)^{-1}\hat{W}'$.

B4. Using the same RSM as for the observed sample, the values of the statistics $ESK_{\mathrm{M}}(v_0)$, $EKU_{\mathrm{M}}(v_0)$, $ESK_{\mathrm{KD}}(v_0)$, $EKU_{\mathrm{KD}}(v_0)$ are calculated from each of these MC samples; in what follows, we will refer to these simulated values as the '*basic simulated statistics*' (BSS): $E^{(j)} = [ESK_{\mathrm{M}}^{(j)}(v_0), EKU_{\mathrm{M}}^{(j)}(v_0), ESK_{\mathrm{KD}}^{(j)}(v_0), EKU_{\mathrm{KD}}^{(j)}(v_0)]'$, $j = 1, \ldots, N$. Using Theorems 3 and 4, it is easy to see that the $N + 1$ vectors $E^{(j)}$, $j = 0, 1, \ldots, N$ are exchangeable under the null hypothesis.

B5. We can then compute a simulated p-value, for any one of the test statistics in $E^{(0)}$: $\hat{p}_N[ESK_{\mathrm{M}}(v_0)]$, $\hat{p}_N[EKU_{\mathrm{M}}(v_0)]$, $\hat{p}_N[ESK_{\mathrm{KD}}(v_0)]$, $\hat{p}_N[EKU_{\mathrm{KD}}(v_0)]$, where $\hat{p}_N[\cdot]$ is defined in equation (27) [for each statistic in E] and can be calculated from the rank of the observed statistic relative to the relevant BSS. The null hypothesis is rejected at level α by the test $ESK_{\mathrm{M}}(v_0)$ if $\hat{p}_N[ESK_{\mathrm{M}}(v_0)] \leq \alpha$, and similarly for the other tests. By the exchangeability of $E^{(j)}$, $j = 0, 1, \ldots, N$, and provided E follows a continuous distribution, this procedure satisfies the size constraint, i.e. $P\big[\hat{p}_N[ESK_{\mathrm{M}}(v_0)] \leq \alpha\big] = \alpha$ under the null hypothesis.

Because the above MC test procedure involves two nested simulations (a first one to get the reference simulated moments, and a second one to get the test statistics), we call it a *double MC test*. The procedure described above allows one to obtain individual simulated p-values for each test statistic. We propose the following combined test statistic, which may be used for all null hypotheses underlying Theorem 4:

$$CSK_{\mathrm{M}}(v_0) = 1 - \min\{\hat{p}_N[ESK_{\mathrm{M}}(v_0)], \hat{p}_N[EKU_{\mathrm{M}}(v_0)]\}, \quad (29)$$
$$CSK_{\mathrm{KD}}(v_0) = 1 - \min\{\hat{p}_N[ESK_{\mathrm{KD}}(v_0)], \hat{p}_N[EKU_{\mathrm{KD}}(v_0)]\}. \quad (30)$$

The intuition here is to reject the null hypothesis if at least one of the individual tests is significant; for convenience, we subtract the minimum p-value from one to obtain a right-sided test. For further reference on these combined tests, see Dufour and Khalaf (2002a), Dufour *et al.* (2003b).

The MC test technique may once again be applied in order to obtain an exact combined test. This can be done by using a *three-stage MC test* (or a *triple MC test*), which involves the estimation of the p-value functions $p_N(\cdot \mid \cdot)$ for individual test statistics, through a preliminary simulation experiment. The algorithm for implementing such a procedure can be described as follows.

C1. Generate a set of reference simulated moments (according to A1–A3), the observed value of $E^{(0)}$ in equation (28), and the N corresponding BSS (following B1–B4).

C2. For each test statistic considered, obtain the p-value functions determined by the BSS (generated at step C1): $p_N(S^{(0)}; S)$, for $S = ESK_{\mathrm{M}}(v_0)$, $EKU_{\mathrm{M}}(v_0)$, $ESK_{\mathrm{KD}}(v_0)$, $EKU_{\mathrm{KD}}(v_0)$, where the function $p_N(S^{(0)}; S)$ is defined in equation (27).

C3. Independent of the previous RSM, BSS and $E^{(0)}$, generate N_1 additional i.i.d. realizations of W according to $\mathscr{F}(v_0)$ in equation (2), and for each draw, compute the pivotal quantities $\hat{W}S_{\hat{W}}^{-1}$ and $T\hat{W}\left(\hat{W}'\hat{W}\right)^{-1}\hat{W}'$. N_1 is chosen so that $\alpha(N_1+1)$ is an integer.

C4. Using the RSM and the N_1 draws generated at steps C1 and C3, compute the corresponding simulated statistics: $EE^{(l)} = [ESK_{\mathrm{M}}^{(l)}(v_0), EKU_{\mathrm{M}}^{(l)}(v_0), ESK_{\mathrm{KD}}^{(l)}(v_0), EKU_{\mathrm{KD}}^{(l)}(v_0)]'$, $l = 1, \ldots, N_1$.

C5. Using the p-value functions $p_N(\cdot\,;\,\cdot)$ obtained at step C2 (and based on the BSS generated at step C1), evaluate the simulated p-values for the observed and the N_1 additional simulated statistics: $\hat{p}_N^{(l)}[S] \equiv p_N(S^{(l)}; S)$, $l = 0, 1, \ldots, N_1$, for $S = ESK_{\mathrm{M}}(v_0)$, $EKU_{\mathrm{M}}(v_0)$, $ESK_{\mathrm{KD}}(v_0)$, $EKU_{\mathrm{KD}}(v_0)$.

C6. From the latter, compute the corresponding values of the combined test statistics:

$$CSK_{\mathrm{M}}^{(l)}(v_0) = 1 - \min\left\{\hat{p}_N[ESK_{\mathrm{M}}^{(l)}(v_0)],\, \hat{p}_N[EKU_{\mathrm{M}}^{(l)}(v_0)]\right\}, \quad l = 0, 1, \ldots, N_1,$$
$$CSK_{\mathrm{KD}}^{(l)}(v_0) = 1 - \min\left\{\hat{p}_N[ESK_{\mathrm{KD}}^{(l)}(v_0)],\, \hat{p}_N[EKU_{\mathrm{KD}}^{(l)}(v_0)]\right\}, \quad l = 0, 1, \ldots, N_1.$$

The vectors $\left(CSK_{\mathrm{M}}^{(l)}(v_0), CSK_{\mathrm{KD}}^{(l)}(v_0)\right)$, $l = 0, 1, \ldots, N_1$, are exchangeable.

C7. The combined test $CSK_{\mathrm{M}}(v_0)$ rejects the null hypothesis at level α if $\hat{p}_{N_1}[CSK_{\mathrm{M}}(v_0)] \equiv p_{N_1}\left(CSK_{\mathrm{M}}^{(0)}; CSK_{\mathrm{M}}(v_0)\right) \leq \alpha$, where $p_{N_1}(\cdot \mid \cdot)$ is based on the simulated variables $CSK_{\mathrm{M}}^{(l)}(v_0), l = 0, 1, \ldots, N_1$; the rule is similar for the test based on $CSK_{\mathrm{KD}}(v_0)$.

The test with critical regions $\hat{p}_{N_1}[CSK_{\mathrm{M}}(v_0)] \leq \alpha$ has level α, because the variables $CSK_{\mathrm{M}}^{(l)}(v_0)$, $l = 0, 1, \ldots, N$, are exchangeable under the null hypothesis. The same holds for the test with critical region $\hat{p}_{N_1}[CSK_{\mathrm{KD}}(v_0)] \leq \alpha$.

We have also studied the following modified version of the omnibus-type tests based on the sum of the skewness and kurtosis statistics:

$$\widetilde{MD} = \begin{bmatrix} SK_{\mathrm{M}} - \overline{SK}_{\mathrm{M}}(v_0) \\ KU_{\mathrm{M}} - \overline{KU}_{\mathrm{M}}(v_0) \end{bmatrix}' \bar{\Delta}_{\mathrm{M}}^{-1}(v_0) \begin{bmatrix} SK_{\mathrm{M}} - \overline{SK}_{\mathrm{M}}(v_0) \\ KU_{\mathrm{M}} - \overline{KU}_{\mathrm{M}}(v_0) \end{bmatrix}, \tag{31}$$

$$\widetilde{JB} = ESK_{\mathrm{KD}}(v_0) + EKU_{\mathrm{KD}}(v_0), \tag{32}$$

where $\bar{\Delta}_{\mathrm{M}}(v_0)$ is a simulation-based estimate of the covariance of SK_{M} and KU_{M}, which can be obtained as outlined in A1–A3; in this case, in addition to the empirical means and standard deviations of simulated SK_{M} and KU_{M} series, we also obtain these empirical covariances. These statistics are obviously less expensive to simulate than the ones based on the smallest p-values. MC p-values can be obtained in a way similar to the one described in B1–B5, except that the matrix $\bar{\Delta}_{\mathrm{M}}(v_0)$ now belongs to the set of moments to be estimated by simulation.

The unknown nuisance parameter case

So far, we have treated the case where the distributional parameter v is known. To account for an unknown v, we obtain a confidence set for this parameter by 'inverting' the above GF tests. Specifically, the confidence set corresponds to the set of v_0 values which are not rejected by the GF test for equation (2) where $v = v_0$ for known v_0. This leads to an estimate of the distributions which best fit the data. This estimate may prove to be very useful for other testing problems regarding the regression under consideration, for it may be easily shown that the usual test statistics for hypothesis on the regression coefficients or error terms will also depend on v (see Beaulieu *et al.*, 2001). When the confidence set for v is empty, the distributional family (2) is rejected.

V. Empirical application

Our empirical analysis focuses on the asset pricing model

$$r_{ti} = a_i + b_i \tilde{r}_{\mathrm{M}_t} + u_{ti}, \quad t = 1, \ldots, T, \; i = 1, \ldots, n, \tag{33}$$

where $r_{ti} = R_{ti} - R_t^F$, $\tilde{r}_{\mathrm{M}_t} = \tilde{R}_{\mathrm{M}_t} - R_t^F$, R_{ti}, $i = 1, \ldots, n$, are returns on n securities for period t, $\tilde{R}_{\mathrm{M}_t}$ are the returns on the market portfolio under consideration, R_t^F is the riskless rate of return, and u_{ti} is a random disturbance. Clearly, this model is a special case of equation (1) where $Y = [r_1, \ldots, r_n]$, $X = [\iota_T, \tilde{r}_M]$, $r_i = (r_{1i}, \ldots, r_{Ti})'$, $\tilde{r}_M = (\tilde{r}_{M1}, \ldots, \tilde{r}_{MT})'$ and u_{ti} are the elements of the matrix U. Our data is obtained from the University of Chicago's Center for Research in Security Prices (CRSP). We use nominal monthly returns over the period, January 1926 to December 1995. We built 12 portfolios of New York Stock Exchange (NYSE) firms using the same classification as Breeden *et al.* (1989): for each month the industry portfolios comprise those firms for which the return, price per common share and number of shares outstanding are recorded by CRSP; portfolios are value-weighted in each month. The market return is measured by the value-weighted NYSE returns, and the risk-free rate by the 1-month Treasury Bill rate. We focus on the multivariate normal and symmetric stable distributions (where the tail thickness parameter is denoted α_s).

Regarding normality tests, Table 1 reveals the following. Although we consider monthly data, normality is definitely rejected except in the last subsample (1990–95). The MC omnibus tests (based on adding up the skewness and kurtosis criteria) and the *min p-value* based combined tests yield the same decision. However, the Mardia-type and the JB-type tests show conflicting decisions in several cases: for the 1941–50, 1956–60, 1966–70, 1971–75 and 1981–85, JB and CSK_{KD} are not significant, whereas our Mardia-type tests are significant; conversely, in 1961–65, although both JB and CSK_{KD} are significant all the Mardia-type tests fail to reject the normal null.

Turning to the stable distributions case, Table 1 reports our test results in the form of confidence sets for the distributional parameters. We observe that Mardia-type confidence sets are the tightest: smaller values of α_s (which signal more extreme kurtosis) are more easily rejected with Mardia-type tests, although, in a few cases, larger values of α_s (which imply tails approaching the normal) are more easily rejected with JB-type tests. The min p-value KD yields tighter confidence sets than its omnibus counterpart. Indeed, $\widetilde{JB}$ suggests that $0.9 \leq \alpha_s \leq 1$ (which signals severely extreme kurtosis) is compatible with our data, whereas all other tests reject $\alpha_s \leq 1$.

These results suggest that it is worthwhile to consider combined JB–Mardia type statistics, such as $CSK_{\mathrm{M/KD}} = 1 - \min\{\hat{p}(CSK_{\mathrm{M}}), \hat{p}(CSK_{\mathrm{KD}})\}$. The flexibility of the MC test method allows one to consider combinations that would be hard to implement otherwise.[5]

[5]See Horswell and Looney (1992) on the cost of combining normality tests. Further details on this empirical study presented here are available in Dufour, Khalaf and Beaulieu (2003a).

TABLE 1

Multivariate normal and symmetric stable distributions: combined skewness–kurtosis-based tests

	Multivariate normal: test p-values					Stable distribution: confidence sets for α_s			
	MSK	CSK_M	$\widetilde{MD}$	*JB*	CSK_{KD}	CSK_M	CSK_{KD}	$\widetilde{MD}$	$\widetilde{JB}$
Sample	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
1927–30	0.001	0.001	0.001	0.007	0.008	1.46–1.92	1.1–1.99	1.36–1.84	0.90–1.99
1931–35	0.001	0.001	0.001	0.001	0.001	1.42–1.90	1.1–1.99	1.38–1.88	0.90–1.99
1936–40	0.001	0.001	0.001	0.020	0.011	1.64–1.98	1.2–1.99	1.56–1.98	0.90–1.99
1941–45	**0.004**	**0.004**	**0.004**	**0.199**	**0.141**	1.66–1.98	1.2–1.99	1.58–1.98	0.90–1.99
1946–50	0.001	0.001	0.001	0.005	0.004	1.58–1.98	1.2–1.99	1.58–1.98	0.90–1.99
1951–55	0.001	0.001	0.001	0.003	0.005	1.64–1.98	1.1–1.99	1.56–1.98	0.90–1.99
1956–60	**0.015**	**0.003**	**0.016**	**0.603**	**0.474**	1.66–1.98	1.2–1.99	1.58–1.98	0.90–1.99
1961–65	**0.736**	**0.631**	**0.151**	**0.008**	**0.029**	1.74–2.00	1.2–2.00	1.66–2.00	0.90–2.00
1966–70	**0.011**	**0.004**	**0.005**	**0.759**	**0.728**	1.66–1.98	1.2–1.99	1.56–1.98	0.90–1.99
1971–75	**0.001**	**0.001**	**0.001**	**0.060**	**0.029**	1.62–1.98	1.2–1.99	1.58–1.98	0.90–1.99
1976–80	0.001	0.001	0.001	0.012	0.030	1.58–1.96	1.1–1.99	1.46–1.96	0.90–1.99
1981–85	**0.001**	**0.001**	**0.002**	**0.305**	**0.154**	1.66–1.98	1.2–1.99	1.58–1.98	0.90–1.99
1986–90	0.024	0.030	0.061	0.009	0.007	1.70–1.98	1.2–1.99	1.64–2.00	0.90–2.00
1991–95	0.917	0.239	0.408	0.127	0.091	1.78–2.00	1.2–2.00	1.78–2.00	0.90–2.00

Notes: Numbers shown in columns (1)–(5) are MC p-values. Columns (1) and (4) refer to MC versions of the original Mardia and JB-type tests (15)–(19). Columns (2) and (5) refer to the min p-value combined skewness/kurtosis criteria (29)–(30). Column (3) reports our dependence-corrected version of Mardia's tests (31). Numbers shown in columns (6)–(9) are values of the kurtosis parameter α_s not rejected by the MC GF tests. Columns (6) and (8) pertain to our combined Mardia-type statistic (29)–(31); columns (7) and (9) are based on the combined Kilian–Demiroglu JB-type statistic (30)–(32). The p-values in bold highlight cases where the various normality tests yield conflicting decisions at the 5% level; in all cases, $N_0 = N = N_1 = 999$.

VI. Conclusion

In this paper, we have proposed a class of exact procedures for testing the GF of the error distribution in possibly non-Gaussian MLR models. In the Gaussian case, our procedures include finite-sample versions of the standard JB and Mardia-type multivariate skewness and kurtosis tests (Mardia, 1970; Kilian and Demiroglu, 2000) as well as new ways of combining skewness and kurtosis measures. The latter are computed on standardized multivariate residuals; so their null distributions do not depend on the unknown error covariance matrix (or the regression coefficients) but remain analytically intractable. The tests are thus implemented using simple, double and triple MC test methods. For non-Gaussian error distributions, we solve the nuisance parameters problem through test 'inversion'.[6]

[6]Due to space limitation, we cannot report the simulation results on the power properties of the procedures discussed here. Such a study is available in Dufour *et al.* (2003a).

The tests suggested were applied to an asset pricing model using monthly returns on NYSE portfolios over 5-year subperiods from 1926 to 1995. We considered testing multivariate normal and stable error distributions. Our results confirm heavy but non-extreme kurtosis, although Mardia- and JB-type tests show conflicting decisions in several cases. The reader may refer to Beaulieu *et al.* (2001) for mean–variance efficiency tests which exploit these results with further focus on elliptical error distributions.

Final Manuscript Received: October 2003

References

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

Beaulieu, M.-C., Dufour, J.-M. and Khalaf, L. (2001). *Testing the CAPM with Possibly Non-Gaussian errors: An Exact Simulation-based Approach*, Technical Report, Département d'économique, Université Laval, and CIREQ, Université de Montréal.

Breeden, D. T., Gibbons, M. and Litzenberger, R. H. (1989). 'Empirical tests of the consumption based CAPM', *Journal of Finance*, Vol. 44, pp. 231–262.

Campbell, Y. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*, Princeton University Press, New Jersey.

D'Agostino, R. B. and Stephens, M. A. (eds) (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Dufour, J.-M. (2002). 'Monte Carlo tests with nuisance parameters: a general approach to finite-sample inference and nonstandard asymptotics in econometrics', *Journal of Econometrics*, forthcoming.

Dufour, J.-M. and Khalaf, L. (2001). 'Monte Carlo test methods in econometrics', in Baltagi B. (ed.), *Companion to Theoretical Econometrics*, Blackwell Companions to Contemporary Economics, Basil Blackwell, Oxford, UK, chapter 23, pp. 494–519.

Dufour, J.-M. and Khalaf, L. (2002a). 'Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions', *Journal of Econometrics*, Vol. 106, pp. 143–170.

Dufour, J.-M. and Khalaf, L. (2002b). 'Simulation based finite and large sample tests in multivariate regressions', *Journal of Econometrics*, Vol. 111, pp. 303–322.

Dufour, J.-M. and Khalaf, L. (2003). 'Finite sample tests in seemingly unrelated regressions', in Giles D. E. A. (ed.), *Computer-Aided Econometrics*, Marcel Dekker, New York, chapter 2, pp. 11–35.

Dufour, J.-M. and Kiviet, J. F. (1998). 'Exact inference methods for first-order autoregressive distributed lag models', *Econometrica*, Vol. 66, pp. 79–104.

Dufour, J.-M., Farhat, A., Gardiol, L. and Khalaf, L. (1998). 'Simulation-based finite sample normality tests in linear regressions', *The Econometrics Journal*, Vol. 1, pp. 154–173.

Dufour, J.-M., Khalaf, L. and Beaulieu, M.-C. (2003a). *Exact Skewness–Kurtosis Tests for Multivariate Normality and Goodness-of-fit in Multivariate Regressions with Application to Asset Pricing Models*, Technical Report, CIRANO and CIREQ, Université de Montréal.

Dufour, J.-M., Khalaf, L., Bernard, J.-T. and Genest, I. (2003b). 'Simulation-based finite-sample tests for heteroskedasticity and ARCH effects', *Journal of Econometrics*, forthcoming.

Dwass, M. (1957). 'Modified randomization tests for nonparametric hypotheses', *Annals of Mathematical Statistics*, Vol. 28, pp. 181–187.

Fiorentini, G., Sentana, E. and Calzolari, G. (2003). *Maximum Likelihood Estimation and Inference in Multivariate Conditionally Heteroskedastic Dynamic Models with Student t Innovations*, Technical Report, Università di Firenze, Florence, Italy.

Gibbons, M. R. (1982). 'Multivariate tests of financial models: a new approach', *Journal of Financial Economics*, Vol. 10, pp. 3–27.

Henze, N. (1994). 'On Mardia's kurtosis test for multivariate normality', *Communications in Statistics, Part A – Theory and Methods*, Vol. 23, pp. 1031–1045.

Holgersson, H. E. T. and Shukur, G. (2001). 'Some aspects of non-normality tests in systems of regression equations', *Communications in Statistics, Simulation and Computation*, Vol. 30, pp. 291–310.

Horswell, R. L. and Looney, S. W. (1992). 'A comparison of tests for multivariate normality that are based on measures of multivariate skewness and kurtosis', *Journal of Statistical Computation and Simulation*, Vol. 42, pp. 21–38.

Horswell, R. L. and Looney, S. W. (1993). 'Diagnostic limitations of skewness coefficients in assessing departures from univariate and multivariate normality', *Communications in Statistics, Part B – Simulation and Computation*, Vol. 22, pp. 437–459.

Jarque, C. M. and Bera, A. K. (1987). 'A test for normality of observations and regression residuals', *International Statistical Review*, Vol. 55, pp. 163–172.

Kiefer, N. M. and Salmon, M. (1983). 'Testing normality in econometric models', *Economic Letters*, Vol. 11, pp. 123–127.

Kilian, L. and Demiroglu, U. (2000). 'Residual-based tests for normality in autoregressions: Asymptotic theory and simulation evidence', *Journal of Business and Economic Statistics*, Vol. 18, pp. 40–50.

Lütkepohl, H. and Theilen, B. (1991). 'Measures of multivariate skewness and kurtosis for tests of nonnormality', *Statistical Papers*, Vol. 32, pp. 179–193.

Mardia, K. V. (1970). 'Measures of multivariate skewness and kurtosis with applications', *Biometrika*, Vol. 57, pp. 519–530.

Mardia, K. V. (1980). 'Tests of univariate and multivariate normality', in Krishnaiah, P. R. (ed.), *Handbook of Statistics 1: Analysis of Variance*, North-Holland, Amsterdam, pp. 279–320.

Richardson, M. and Smith, T. (1993). 'A test for multivariate normality in stock returns', *Journal of Business*, Vol. 66, pp. 295–321.

Shanken, J. (1996). 'Statistical methods in tests of portfolio efficiency: A synthesis', in Maddala, G. S. and Rao, C. R. (eds), *Handbook of Statistics 14: Statistical Methods in Finance*, North-Holland, Amsterdam, pp. 693–711.

Stewart, K. G. (1997). 'Exact testing in multivariate regression', *Econometric Reviews*, Vol. 16, pp. 321–352.

Thode, H. C. Jr. (2002). Testing for Normality, number 164 in *Statistics: Textbooks and Monographs*, Marcel Dekker, New York.

Zhou, G. (1993). 'Asset-pricing tests under alternative distributions', *The Journal of Finance*, Vol. 48, pp. 1927–1942.