Finite-sample multivariate tests of asset pricing models with coskewness *

Marie-Claude Beaulieu † Jean-Marie Dufour ‡ Lynda Khalaf §
Université Laval McGill University Carleton University

First version: July 2002
Revised: July 2002, April 2004, April 2005, September 2007
This version: February 2008
Compiled: March 16, 2008, 9:44pm

This paper is forthcoming in the *Computational Statistics and Data Analysis*.

* The authors thank the Editor David Belsley, two anonymous referees, Raja Velu, Craig MacKinlay, participants at the 2005 Finite Sample Inference in Finance II, the 2006 European Meetings of the Econometric Society, the 2006 New York Econometrics Camp, and the 2007 International workshop on Computational and Financial Econometrics conferences, for several useful comments. This work was supported by the William Dow Chair in Political Economy (McGill University), the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (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 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), the Fonds de recherche sur la nature et les technologies (Québec), the Chaire RBC en innovations financières (Université Laval), and NATECH (Government of Québec).

† RBC Chair in Financial Innovations and Département de finance et assurance, Université Laval, CIRANO, and Centre Interuniversitaire sur le risque, les politiques economiques et l'emploi (CIRPEE). Mailing address: Département de finance et assurance, Pavillon Palasis-Prince, Université Laval, Québec, Québec G1K 7P4, Canada. TEL: 1 (418) 656-2926, FAX: 1 (418) 656-2624; e-mail: Marie-Claude.Beaulieu@fas.ulaval.ca

‡ William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

§ Canada Research Chair in Environmental and Financial Econometric Analysis (Université Laval), Economics Department, Carleton University, CIREQ, and Groupe de recherche en économie de l'énergie, de l'environnement et des ressources naturelles (GREEN), Université Laval. Mailing address: Economics Department, Carleton University, Loeb Building 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada. TEL: 1 (613) 520 2600 ext. 8697; FAX: 1 (613) 520 3906; e-mail: Lynda_Khalaf@carleton.ca

ABSTRACT

Finite-sample inference methods are proposed for asset pricing models with unobservable risk-free rates and coskewness, specifically, for the *quadratic market model* (QMM) which incorporates the effect of asymmetry of return distribution on asset valuation. In this context, exact tests are appealing for several reasons: (i) the increasing popularity of such models in finance, (ii) the fact that traditional market models (which assume that asset returns move proportionally to the market) have not fared well in empirical tests, (iii) finite-sample tests for the QMM are unavailable even with Gaussian errors. Empirical models are considered where the procedure to assess the significance of coskewness preference is LR-based, and relates to the statistical and econometric literature on dimensionality tests which are interesting in their own right. Exact versions of these tests are obtained, allowing for non-normality of fundamentals. A simulation study documents the size and power properties of asymptotic and finite-sample tests. Empirical results with well known data sets reveal temporal instabilities over the full sampling period, namely 1961-2000, though tests fail to reject the QMM restrictions over 5-year subperiods.

Key words: capital asset pricing model; CAPM; quadratic market model; QMM; Black; mean-variance efficiency; non-normality; weak identification; multivariate linear regression; uniform linear hypothesis; exact test; Monte Carlo test; bootstrap; nuisance parameters; specification test; diagnostics; GARCH; variance ratio test.

Journal of Economic Literature classification: C3; C12; C33; C15; G1; G12; G14.

1. Introduction

The traditional market model, which supposes that expected asset returns move proportionally to their market beta risk, has not fared well in empirical tests; see *e.g.* Campbell (2003). These failures have spurred studies on possible nonlinearities and asymmetries in the dependence between asset and market returns. In particular, the quadratic market model (QMM) was proposed to extend the standard CAPM framework to incorporate the effect of return distribution asymmetry on asset valuation; see Kraus and Litzenberger (1976). A number of well known empirical studies [e.g. Dittmar (2002), Harvey and Siddique (2000), Simaan (1993), Barone-Adesi (1985), Barone-Adesi, Gagliardini and Urga (2004*a*, 2004*b*), Smith (2007)] have attempted to assess such asset pricing models; yet these studies are only asymptotically justified. Exact QMM tests seem lacking even with Gaussian errors.

We consider empirical settings based on multivariate linear regressions (MLR), as in Barone-Adesi (1985) and Barone-Adesi, Gagliardini and Urga (2004*a*, 2004*b*). These studies model coskewness preference via an Arbitrage Pricing Theory (APT) framework, leading to nonlinear cross-equation constraints on an MLR-based QMM. The associated MLR includes the expected return in excess of the zero-beta portfolio (assumed unknown and denoted γ) and an extra regressor: the expected excess returns on a portfolio perfectly correlated with the squared market returns; the latter parameter (denoted θ) is also unobservable and should be estimated in empirical applications. This formulation is interesting because it nests Black's fundamental model [Black (1972), Gibbons (1982), Shanken (1996)] theoretically and statistically.

Empirical tests of several well known asset pricing models [see Shanken (1996), Campbell, Lo and MacKinlay (1997), Dufour and Khalaf (2002) and Beaulieu, Dufour and Khalaf (2007)] are often conducted within the MLR framework. In this context however, and as may be checked from the above-cited references, discrepancies between asymptotic and finite-sample distributions are documented and are usually ascribed to the curse of dimensionality: as the number of equations increases, the number of error cross-correlations grows rapidly which leads to reductions in degrees of freedom and to test size distortions. In the QMM case, the model further involves nonlinear restrictions whose identification raises serious non-regularities: γ and θ are not identified over the whole parameter space, which can strongly affect the distributions of estimators and test statistics, leading to the failure of standard asymptotics. For references on weak-identification problems, see *e.g.* Dufour (1997, 2003), Stock, Wright and Yogo (2002), Dufour and Taamouti (2005, 2007), Joseph and Kiviet (2005), Khalaf and Kichian (2005), Kiviet and Niemczyk (2007), Bolduc, Khalaf and Moyneur (2008), and the references therein. We propose here exact likelihood ratio (LR) type tests immune to such difficulties.

To obtain an exact version of the QMM test under consideration, we formulate the problem as a dimensionality test on the parameters of a MLR, and derive closed-form analytical expressions for estimates and test statistics. Exact dimensionality tests are available under Gaussian fundamentals [see *e.g.* Zhou (1991), Zhou (1995) and Velu and Zhou (1999) who point out references to the statistical literature], yet these procedures have not been applied to the QMM framework. Of course, empirical evidence on non-normalities of financial returns may lead one to question the usefulness of Gaussian based exact tests. The results of Beaulieu, Dufour and Khalaf (2005, 2006, 2007) related to the standard CAPM confirm this issue: despite relaxing normality (using provably exact non-Gaussian tests), the linear CAPM is still rejected for several subperiods. We thus propose to extend the statistical framework of Beaulieu et al. (2007) to the QMM case.

The paper is organized as follows. Section 2 sets the framework. Section 3 presents estimates and test statistics. Distributional results are discussed in sections 4 and 5. In Section 6, we study the problem of testing the homogeneity hypothesis. In section 7, we report a simulation study which assesses the finite-sample performance of available asymptotic tests as well as the power properties of our proposed exact tests. Section 8 reports our empirical results and section 9 concludes.

2. Framework

Let R_{it}, $i = 1, \ldots, n$, be returns on n securities for period t, $t = 1, \ldots, T$ and $\tilde{R}_{\mathrm{M}t}$ the returns on a market portfolio under consideration. Kraus and Litzenberger (1976)'s quadratic market model (QMM) takes the following form:

$$R_{it} = a_i + b_i \tilde{R}_{\mathrm{M}t} + c_i \tilde{R}_{\mathrm{M}t}^2 + u_{it}\,, \quad i = 1, \ldots, n,\ t = 1, \ldots, T, \tag{2.1}$$

where the u_{it} are random disturbances. Our main statistical results require that the vectors $V_t = (u_{1t}, \ldots, u_{nt})'$, $t = 1, \ldots, T$, satisfy the following distributional assumption:

$$V_t = JW_t\,,\ t = 1,\ \ldots\ ,\ T\,, \tag{2.2}$$

where J is an unknown nonsingular $n \times n$ matrix and the distribution of $w = vec(W)$ with $W = [W_1, \ldots, W_T]'$ is: either (i) fully specified, or (ii) specified up to a nuisance parameter ν. We will treat first the case where ν is known; the case of unknown ν is discussed in section 5. We also study the special cases where the errors are *i.i.d.* following: (i) a multivariate Gaussian distribution, *i.e.*,

$$W_1,\ \ldots\ ,\ W_T \sim N[0,\ I_n] \tag{2.3}$$

or (ii) a multivariate Student-t distribution with degrees-of-freedom κ, *i.e.*,

$$W_t = Z_{1t}/(Z_{2t}/\kappa)^{1/2}, \tag{2.4}$$

where $Z_{1t} \sim N[0,\ I_n]$ and Z_{2t} is a $\chi^2(\kappa)$ variate independent from Z_{1t}.

Using Ross' Arbitrage Pricing Theory, Barone-Adesi (1985) derived a model of equilibrium based on (2.1), which entails the following restrictions:

$$\mathcal{H}_Q : a_i + \gamma(b_i - 1) + c_i\theta = 0\,, \quad i = 1,\ \ldots\ ,\ n,\ \text{for some } \gamma \text{ and } \theta. \tag{2.5}$$

$\mathcal{H}_Q$ is nonlinear since γ and θ are unknown. Clearly, γ may be weakly identified if the b_i coefficients are close to 1, and θ may be weakly identified if the c_i coefficients are close to 0.

For any specified values γ_0 and θ_0, the hypothesis

$$\mathcal{H}_Q(\gamma_0, \theta_0) : a_i + \gamma_0(b_i - 1) + c_i\theta_0 = 0\,,\ i = 1,\ \ldots\ ,\ n, \tag{2.6}$$

is linear. This observation underlies our exact bound test procedure; we can also use pivotal statistics associated with the latter hypothesis as in Beaulieu, Dufour and Khalaf (2006) to obtain a joint confidence set for γ and θ.

To simplify the presentation, let us transform the model as follows:

$$R_{it} - \tilde{R}_{\mathrm{M}t} = a_i + (b_i - 1)\tilde{R}_{\mathrm{M}t} + c_i \tilde{R}^2_{\mathrm{M}t} + u_{it}\,,\ t = 1, \ldots, T,\ i = 1, \ldots, n. \tag{2.7}$$

The above model is a special case of the following MLR:

$$Y = XB + U, \tag{2.8}$$

where $Y = [Y_1, \ldots, Y_n]$ is $T \times n$ matrix, X is $T \times k$ matrix with rank k and is assumed fixed, and $U = [U_1, \ldots, U_n] = [V_1, \ldots, V_T]'$ is the $T \times n$ matrix of error terms. In most cases of interest, we also have $n \geq k$. Clearly, (2.7) corresponds to the case where:

$$Y = [R_1 - \tilde{R}_{\mathrm{M}}, \ldots, R_n - \tilde{R}_{\mathrm{M}}]\,, \quad X = [\iota_T, \tilde{R}_{\mathrm{M}}, \tilde{R}^2_{\mathrm{M}}], \tag{2.9}$$

$$R_i = (R_{1i}, \ldots, R_{Ti})', \quad \tilde{R}_{\mathrm{M}} = (\tilde{R}_{1\mathrm{M}}, \ldots, \tilde{R}_{T\mathrm{M}})'\,, \quad \tilde{R}^2_{\mathrm{M}} = (\tilde{R}^2_{1\mathrm{M}}, \ldots, \tilde{R}^2_{T\mathrm{M}})'\,, \tag{2.10}$$

$$B = \begin{bmatrix} a_1 & \cdots & a_n \\ b_1 - 1 & \cdots & b_n - 1 \\ c_1 & \cdots & c_n \end{bmatrix}. \tag{2.11}$$

In this context, the (Gaussian) quasi likelihood ratio (QLR) criterion for testing (2.5) is:

$$LR_Q = T \ln(|\hat{\Sigma}_Q| / |\hat{\Sigma}|) \tag{2.12}$$

where $\hat{\Sigma}$ is the unrestricted (Gaussian) quasi maximum likelihood estimator (QMLE) of Σ and $\hat{\Sigma}_Q$ is the restricted QMLE of Σ under $\mathcal{H}_Q$. Note that

$$\hat{\Sigma} = \hat{U}'\hat{U}/T = Y'MY/T\,,\ \hat{U} = Y - X\hat{B}\,,\ \hat{B} = (X'X)^{-1}X'Y,\ M = I - X(X'X)^{-1}X'. \tag{2.13}$$

Conformably, let $\hat{a}_i$, $\hat{b}_i$ and $\hat{c}_i$ refer to the QMLEs of a_i, b_i and c_i. Barone-Adesi (1985) implements a linearization procedure as in Gibbons (1982) to approximate the latter statistic; in a related framework which corresponds to a known γ, Barone-Adesi, Gagliardini and Urga (2004*b*) obtains a test statistic of this form using iterative numerical maximization. In what follows, we propose: (i) simple eigenvalue-based non-iterative formula to derive (2.12), (ii) exact bounds on p-values and, (iii) bootstrap-type cut-off p-values.

Our analysis reveals similarities with the statistical foundations of MLR-based tests of Black's version of the CAPM. Indeed, if $c_i = 0,\ i = 1,\ \ldots\ ,\ n$, the above model nests Black's model. The statistical results we provide here thus extend Shanken (1986), Zhou (1991) and Velu and Zhou (1999) to the three-moment CAPM case. Some results from Zhou (1995) are also relevant for the Gaussian case, although Zhou (1995) did not study the three-moment CAPM.

Alternative formulations of the three moment CAPM allow for a less restrictive nonlinear hypothesis [see Barone-Adesi, Gagliardini and Urga (2004*a*)] as follows:

$$\bar{\mathcal{H}}_Q : a_i + \gamma(b_i - 1) + c_i\theta = \phi,\ i = 1,\ \ldots\ ,\ n, \quad \text{for some } \gamma,\ \theta \text{ and } \phi, \tag{2.14}$$

where ϕ is the same across portfolios. We will consider this hypothesis in section 6.

3. Constrained estimation and test statistics

In this section, we provide convenient non-iterative formulae for computing QMLE-based test statistics for $\mathcal{H}_Q(\gamma_0,\ \theta_0)$ and $\mathcal{H}_Q$.

3.1. Linear case

Hypothesis $\mathcal{H}_Q(\gamma_0,\ \theta_0)$ in (2.6) is a special case of

$$\mathcal{H}(C_0) : C_0 B = 0, \tag{3.1}$$

where $C_0 \in \mathcal{M}(k-r,\ k)$, $0 \le k-r \le k$ and $\mathcal{M}(m_1,\ m_2)$ denotes the set of full-rank $m_1 \times m_2$ matrices with real elements. In this case, the constrained QMLEs are:

$$\hat{B}(C_0) \;=\; \hat{B} - \left(X'X\right)^{-1} C_0' \left[C_0 \left(X'X\right)^{-1} C_0'\right]^{-1} C_0 \hat{B}\,, \tag{3.2}$$

$$\hat{\Sigma}(C_0) = \hat{\Sigma} + \hat{B}'C_0'\big[C_0\left(X'X\right)^{-1}C_0'\big]^{-1}C_0\hat{B} = Y'M(C_0)Y\,, \tag{3.3}$$

$$M(C_0) = M + X(X'X)^{-1}C_0'[C_0(X'X)^{-1}C_0']^{-1}C_0(X'X)^{-1}X', \tag{3.4}$$

where M is defined in (2.13). Furthermore, the standard LR and Wald statistics for testing $\mathcal{H}(C_0)$ – denoted respectively $LR(C_0)$ and $\mathcal{W}(C_0)$ – can be expressed in terms of the eigenvalues $\hat{\mu}_1(C_0) \geq \hat{\mu}_2(C_0) \geq \cdots \geq \hat{\mu}_n(C_0)$ of the matrix $\hat{\Sigma}(C_0)^{-1}[\hat{\Sigma}(C_0) - \hat{\Sigma}]$:

$$LR(C_0) = -T\ln\left(|I - \hat{\Sigma}(C_0)^{-1}[\hat{\Sigma}(C_0) - \hat{\Sigma}]|\right) = -T\sum_{i=1}^{l}\ln[1 - \hat{\mu}_i(C_0)]\,, \tag{3.5}$$

$$\mathcal{W}(C_0) = T\,\mathrm{trace}\left(\hat{\Sigma}^{-1}[\hat{\Sigma}(C_0) - \hat{\Sigma}]\right) = T\sum_{i=1}^{l}\frac{\hat{\mu}_i(C_0)}{1 - \hat{\mu}_i(C_0)}\,, \tag{3.6}$$

where $l \equiv \min\{k - r,\, n\}$ is the rank of $\hat{\Sigma}(C_0)^{-1}[\hat{\Sigma}(C_0) - \hat{\Sigma}]$. If $k \leq n$, we have $l = k - r$. Applying these expressions to test $\mathcal{H}_Q(\gamma_0, \theta_0)$ leads to the following constrained QMLEs:

$$\hat{B}(\gamma_0,\, \theta_0) = \hat{B}(C_0), \qquad \hat{\Sigma}(\gamma_0,\, \theta_0) = \hat{\Sigma}(C_0), \qquad M(\gamma_0,\, \theta_0) = M(C_0), \tag{3.7}$$

where $\hat{B}(C_0)$, $\hat{\Sigma}(C_0)$ and $M(C_0)$ obtain from (3.2) - (3.4) with X and Y as defined in (2.9) and $C_0 = [1,\, \gamma_0,\, \theta_0]$. In this case, $l = \mathrm{rank}(C_0) = k - r = 1$, so that the matrix $\hat{\Sigma}(\gamma_0,\, \theta_0)^{-1}[\hat{\Sigma}(\gamma_0,\, \theta_0) - \hat{\Sigma}]$ has only one non-zero root which we denote $\hat{\mu}(\gamma_0,\, \theta_0)$, and the LR and Wald test statistics are thus monotonic transformations of each other given by:

$$LR(\gamma_0,\, \theta_0) = -T\ln[1 - \hat{\mu}(\gamma_0,\, \theta_0)]\,, \qquad \mathcal{W}(\gamma_0,\, \theta_0) = T\left(\frac{\hat{\mu}(\gamma_0,\, \theta_0)}{1 - \hat{\mu}(\gamma_0,\, \theta_0)}\right)\,. \tag{3.8}$$

3.2. Nonlinear case

For tests on the rank of a matrix, Gouriéroux, Monfort and Renault (1993, 1995) provide the following formulae to test hypotheses of the form

$$\mathcal{H}_{NL} : CB = 0\,,\ \text{for some } C \in \mathcal{M}(k - r,\, k), \tag{3.9}$$

in the context of the MLR (2.8). The LR and Wald statistics may be written as:

$$LR_{_{NL}} = \min_{C \in \mathcal{M}(k-r,\,k)} LR(C) = -T \sum_{i=r+1}^{k} \ln(1-\hat{\lambda}_i), \tag{3.10}$$

$$\mathcal{W}_{_{NL}} = \min_{C \in \mathcal{M}(k-r,\,k)} \mathcal{W}(C) = T \sum_{i=r+1}^{k} \frac{\hat{\lambda}_i}{(1-\hat{\lambda}_i)}, \tag{3.11}$$

where $LR(C)$, $\mathcal{W}(C)$ are as defined in (3.5)-(3.6) and $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_k$ are the eigenvalues of

$$R_{XY} = (X'X)^{-1}X'Y(Y'Y)^{-1}Y'X. \tag{3.12}$$

Let $\hat{e}_1,\, \hat{e}_2,\, \ldots\,,\, \hat{e}_k$ denote the eigenvectors associated with $\hat{\lambda}_1,\, \hat{\lambda}_2,\, \ldots\,,\, \hat{\lambda}_k$, normalized so that

$$[\hat{e}_{r+1},\, \ldots\,,\, \hat{e}_k]'\,(X'X)\,[\hat{e}_{r+1},\, \ldots\,,\, \hat{e}_k] = I. \tag{3.13}$$

Then $\hat{C}_{NL} = [\hat{e}_{r+1},\, \ldots\,,\, \hat{e}_k]'\,(X'X)$ gives the QMLE of C. Note that $\mathcal{H}_{NL}$ is equivalent to $B = \delta\beta$ for some δ, where β is $r \times n$ and δ is a $k \times r$ matrix of rank r which are linked to C through the condition

$$C\delta = 0. \tag{3.14}$$

In the same vein, when δ_0 is a known $k \times r$ matrix of rank r such that $C\delta_0 = 0$, then

$$\mathcal{H}(C_0) \Leftrightarrow B = \delta_0\beta\,. \tag{3.15}$$

Consider the transformed regression $Y = \hat{X}(\hat{\delta})\beta + U$ and associated OLS estimator $\hat{\beta}(\hat{\delta})$ where

$$\hat{\delta} = [\hat{e}_1,\, \ldots\,,\, \hat{e}_r]\,, \quad \hat{X}(\hat{\delta}) = X\hat{\delta}, \quad \hat{\beta}(\hat{\delta}) = \left[\hat{X}(\hat{\delta})'\hat{X}(\hat{\delta})\right]^{-1}\hat{X}(\hat{\delta})'Y. \tag{3.16}$$

Then $\hat{\delta}$ provides a QMLE for δ, and QMLEs of B and Σ (denoted $\hat{B}_{NL}$ and $\hat{\Sigma}_{NL}$) obtain as

$$\hat{B}_{NL} = \hat{\delta}\hat{\beta}(\hat{\delta})\,, \quad \hat{\Sigma}_{NL} = \hat{U}(\hat{\delta})'\hat{U}(\hat{\delta})/T, \quad \hat{U}(\hat{\delta}) = Y - \hat{X}\hat{\beta}(\hat{\delta})\,. \tag{3.17}$$

Under standard regularity assumptions, $LR_{_{NL}} \overset{asy}{\sim} \chi^2\big((n-r)(k-r)\big)$ and $\mathcal{W}_{_{NL}} \overset{asy}{\sim} \chi^2\big((n-r)(k-r)\big)$; see Gouriéroux, Monfort and Renault (1993, 1995).

The QMM restrictions $\mathcal{H}_Q$ in (2.5) are a special case of (3.9) with $C = [1,\, \gamma,\, \theta]$. Since $k = 3$ and $k-r = 1$, we have $r = 2$. The fact that one element of C is equal to one implies a normalization that does not alter the formulae for the test statistics. Let $\hat{\lambda}_{Q1} \geq \hat{\lambda}_{Q2} \geq \hat{\lambda}_{3Q}$ and $\hat{e}_{Q1},\, \hat{e}_{Q2},\, \hat{e}_{Q3}$ denote the the eigenvalues of (3.12) [there are $k = 3$ non-zero roots] and associated eigenvectors [normalized as in (3.13)] where X and Y are as in (2.9), and define

$$\hat{\delta}_Q = [\hat{e}_{Q1},\, \hat{e}_{Qr}]\,. \tag{3.18}$$

Then the associated LR and Wald statistics are monotonic transformations of each other given by:

$$LR_Q = \inf_{\gamma,\,\theta}\{LR(\gamma,\,\theta)\} = -T\ln(1-\hat{\lambda}_{Q3})\,, \qquad \mathcal{W}_Q = \inf_{\gamma,\,\theta}\{\mathcal{W}(\gamma,\theta)\} = T\frac{\hat{\lambda}_{Q3}}{(1-\hat{\lambda}_{Q3})}\,, \tag{3.19}$$

where $LR(\gamma,\,\theta)$ and $\mathcal{W}(\gamma,\theta)$ are as defined in (3.8). By (3.17), constrained QMLEs can be written as:

$$\hat{B}_Q = \hat{\delta}_Q\hat{\beta}\big(\hat{\delta}_Q\big)\,, \quad \hat{\Sigma}_Q = \hat{U}\big(\hat{\delta}_Q\big)'\hat{U}\big(\hat{\delta}_Q\big)/T, \quad \hat{U}\big(\hat{\delta}_Q\big) = Y - \hat{X}\hat{\beta}\big(\hat{\delta}_Q\big), \tag{3.20}$$

where $\hat{X}(\hat{\delta})$ and $\hat{\beta}(\hat{\delta})$ correspond to (3.16) replacing $\hat{\delta}$ with $\hat{\delta}_Q$. QMLEs for γ and θ [denoted $\hat{\gamma}_Q$ and $\hat{\theta}_Q$] may be obtained from (3.18) using the orthogonality condition (3.14). To the best of our knowledge, the latter formulae for Barone-Adesi (1985)'s model have not been applied to date. The original application in Barone-Adesi (1985) relied on a linearized version of the model and a pre-set estimator for θ which is not the QMLE; while Barone-Adesi et al. (2004*a*, 2004*b*) do tackle the nonlinear problem, they follow a standard iterative based numerical MLE.

Under strong-identification regularity assumptions, $LR_{_Q} \overset{asy}{\sim} \chi^2\,(n-2)$ and $\mathcal{W}_{_Q} \overset{asy}{\sim} \chi^2(n-2)$. However, it is well known that the latter approximations: (i) perform poorly in finite samples, particularly if n is large relative to T, and (ii) may lead to severe size distortions, because the underlying asymptotics are not corrected for weak-identification; see Dufour and Khalaf (2002) and Dufour (1997, 2003).

4. Exact distributional results

In Dufour and Khalaf (2002), we have recently derived several exact distributional results regarding the QLR criteria discussed in the previous section for linear and nonlinear hypotheses. The following results are relevant to the problem under consideration.

4.1. Tests on QMM parameters

Let us first consider results relevant to the hypothesis (2.6). This result is a special case of Theorem 3.1 in Dufour and Khalaf (2002).

Theorem 4.1 DISTRIBUTION OF TESTS FOR UNIFORM LINEAR HYPOTHESES. *Under* (2.2), (2.8) *and* (3.1), $LR(C_0)$ *and* $\mathcal{W}(C_0)$ [*defined in* (3.5)-(3.6)] *are respectively distributed like*

$$\overline{LR}(C_0) = -T\sum_{i=1}^{l} \ln[1-\mu_i(C_0)]\,, \qquad \overline{\mathcal{W}}(C_0) = T\sum_{i=1}^{l} \frac{\mu_i(C_0)}{1-\mu_i(C_0)}, \tag{4.1}$$

where $\mu_1(C_0) \geq \mu_2(C_0) \geq \cdots \geq \mu_l(C_0)$ *are the eigenvalues of the matrix*

$$\bar{F}_L(C_0) \equiv [W'M(C_0)W]^{-1}[W'M(C_0)W - W'MW], \tag{4.2}$$

M *and* $M(C_0)$ *are defined in* (2.13) *and* (3.4) *and* $W = [W_1, \ldots, W_T]'$ *is defined by* (2.2).

In view of (3.15), the constrained projection matrix can also be calculated as

$$M(C_0) = M(\delta_0) = I - \hat{X}(\delta_0)(\hat{X}(\delta_0)'\hat{X}(\delta_0))^{-1}\hat{X}(\delta_0)\,, \quad \hat{X}(\delta_0) = X\delta_0. \tag{4.3}$$

For certain values of $k-r$ and normal errors, the null distribution in question reduces to the F distribution. For instance, if $k-r=1$, then

$$\frac{T-(k-1)-n}{n}\left[\left(|\hat{\Sigma}(C_0)|/|\hat{\Sigma}|\right)-1\right] \sim F\big(n,\, T-(k-1)-n\big)\,. \tag{4.4}$$

These results may be applied to the problem at hand for inference on γ and θ. It will be convenient to study first the problem of testing hypotheses of the form $\mathcal{H}_Q(\gamma_0, \theta_0)$.

Theorem 4.2 DISTRIBUTION OF LINEAR QMM TEST STATISTICS. *Under* (2.2) *and* (2.6) - (2.9), $LR(\gamma_0, \theta_0)$ *and* $\mathcal{W}(\gamma_0, \theta_0)$ [*defined in* (3.8)] *are respectively distributed like*

$$\overline{LR}(\gamma_0, \theta_0) = -T \ln[1 - \mu(\gamma_0, \theta_0)], \qquad \overline{\mathcal{W}}(\gamma_0, \theta_0) = T \frac{\mu(\gamma_0, \theta_0)}{1 - \mu(\gamma_0, \theta_0)}, \tag{4.5}$$

where $\mu(\gamma_0, \theta_0)$ *is the non-zero eigenvalue of the matrix*

$$\bar{F}_L(\gamma_0, \theta_0) \equiv [W'M(\gamma_0, \theta_0)W]^{-1}[W'M(\gamma_0, \theta_0)W - W'MW], \tag{4.6}$$

with M, $M(\gamma_0, \theta_0)$ *and* W *as defined in* (2.13), (3.7) *and* (2.2), *respectively.*

In this case, the tests based on $LR(\gamma_0, \theta_0)$ and $\mathcal{W}(\gamma_0, \theta_0)$ are equivalent because they are monotonic transformations of each other. The latter theorem shows that the exact distributions of $LR(\gamma_0, \theta_0)$ and $\mathcal{W}(\gamma_0, \theta_0)$ do not depend on any unknown nuisance parameter as soon as the distribution of W is completely specified. In particular, the values of B and J are irrelevant, though it depends in general upon X, γ_0 and θ_0. Although possibly non standard, the relevant distributions may easily be simulated. For generality and further reference, we present a generic algorithm to obtain a MC p-value based on a pivotal statistic of the form $S(y, X)$, that can be written as a pivotal function of W (in (2.2)) knowing X, formally

$$S(y, X) = \overline{S}\,(W, X)\,, \tag{4.7}$$

where the distribution underlying W is fully specified or is specified up to the parameter ν, in which case the following conditions on ν.

1. Let S_0 denote the test statistic calculated from the observed data set.

2. For a given number N of replications, draw $W^j = [W_1^j, \ldots, W_n^j]$, $j = 1, \ldots, N$, as in (2.2), conditional on ν when relevant. Correspondingly obtain $S^j = \overline{S}(W^j, X)$, $j =$

$1\ ,\ \ldots\ ,\ N$. For instance, in the case of the QLR statistic underlying Theorem (**4.1**), use the pivotal expression $\overline{LR}(\gamma_0,\, \theta_0)$ as defined by (4.5)-(4.6) for $\overline{S}\,(W,\ X)$.

3. Given the series of simulated statistics $S_1,\ ...,\ S_N$, compute

$$\hat{p}_N^{PMC}(S_0) = \frac{N\hat{G}_N(S_0\ ) + 1}{N+1}, \tag{4.8}$$

where $N\hat{G}_N(S_0)$ is the number of simulated values which are greater than or equal to S_0. If step 2 conditions on ν, we propose to modify notation to emphasize this fact, so $\hat{p}_N^{PMC}(S_0)$ is replaced by $\hat{p}_N^{PMC}(S_0|\nu)$. We use the superscript "PMC" to emphasize that the p-value is based on a proper pivot.

4. The MC critical region is

$$\hat{p}_N^{PMC}(S_0) \leq \alpha, \quad 0 < \alpha < 1. \tag{4.9}$$

If $\alpha(N+1)$ is an integer, then under the null hypothesis, $\mathsf{P}\left[\hat{p}_N^{PMC}(S_0) \leq \alpha\right] = \alpha$ and $\mathsf{P}\left[\hat{p}_N^{PMC}(S_0|\nu) \leq \alpha\right] = \alpha$ for known ν.

The latter algorithm applies to the statistic $LR(\gamma_0,\, \theta_0)$ [we focus on the latter rather than on $\mathcal{W}(\gamma_0,\, \theta_0)$ to compare our results with published works and in view of their functional equivalence] using (4.5). For further reference, we call the MC p-value so obtained

$$p_N^{PMC}[\gamma_0,\, \theta_0, \nu] \equiv \hat{p}_N^{PMC}[LR(\gamma_0,\, \theta_0)|\nu] \tag{4.10}$$

where N is the number of MC replications, the superscript PMC emphasizes the fact that the simulated statistic is a proper pivot and ν refers to the parameter of the error distribution.

Since the procedure just described allows one to test any hypothesis of the form $\mathcal{H}_Q(\gamma_0,\, \theta_0)$, which sets the values of both γ and θ, we can build a confidence set for $(\gamma,\, \theta)$, by considering all pairs $(\gamma_0,\, \theta_0)$ which are not rejected, *i.e.* such that $\hat{p}_N^{PMC}[LR(\gamma_0,\, \theta_0)|\nu] > \alpha$:

$$\mathcal{C}(\alpha;\, \gamma,\, \theta) = \left\{(\gamma_0,\, \theta_0) : \hat{p}_N^{PMC}[LR(\gamma_0,\, \theta_0)|\nu] > \alpha\right\}. \tag{4.11}$$

By construction, we have:

$$\mathsf{P}[(\gamma,\,\theta) \in \mathcal{C}(\alpha;\,\gamma,\,\theta)] \geq 1 - \alpha\,. \tag{4.12}$$

In the special case where the errors are *i.i.d.* Gaussian as in (2.3), the null distribution of $LR(\gamma_0,\,\theta_0)$ takes a simpler form. Since $\mathrm{rank}(C_0) = k - r = 1$, we have:

$$\frac{(T-2-n)}{n}\big[\big(|\hat{\Sigma}(\gamma_0,\,\theta_0)|/|\hat{\Sigma}|\big) - 1\big] \sim F(n,\,T-2-n\,)\,. \tag{4.13}$$

In this case, the distribution of $LR(\gamma_0,\,\theta_0)$ does not depend on X, γ_0 or θ_0. This convenient feature is, however, specific to the Gaussian error distribution (though we cannot exclude the possibility that it holds for other distributions).

4.2. Tests for QMM

Let us now consider the problem of testing $\mathcal{H}_Q$ in (2.5). We propose a number of possible avenues in order to deal with the nonlinear nature of the problem.

For that purpose, we start from the confidence set $\mathcal{C}(\alpha;\,\gamma,\,\theta)$ for $(\gamma,\,\theta)$. Under the assumption $\mathcal{H}_Q$, $\mathcal{C}(\alpha;\,\gamma,\,\theta)$ contains the true parameter vector $(\gamma,\,\theta)$ with probability at least $1-\alpha$, in which case the set $\mathcal{C}(\alpha;\,\gamma,\,\theta)$ is obviously not empty. By contrast, if $\mathcal{H}_Q$ does not hold, no vector $(\gamma_0,\,\theta_0)$ does satisfy $\mathcal{H}_Q$: if the test is powerful enough, all proposed values will be rejected, in which case $\mathcal{C}(\alpha;\,\gamma,\,\theta)$ is empty. The size condition (4.12) entails the following property: under $\mathcal{H}_Q$,

$$\mathsf{P}[\mathcal{C}(\alpha;\,\gamma,\,\theta) = \emptyset] \leq \alpha\,. \tag{4.14}$$

Thus, by checking whether we have $\mathcal{C}(\alpha;\,\gamma,\,\theta) = \emptyset$, *i.e.*

$$\{(\gamma_0,\,\theta_0) : \hat{p}_N^{PMC}[LR(\gamma_0,\,\theta_0)|\nu] > \alpha\} = \emptyset\,, \tag{4.15}$$

we get a critical region with level α for $\mathcal{H}_Q$.

In view of the fact that the set $\mathcal{C}(\alpha;\,\gamma,\,\theta)$ only involves two parameters, it can be established fairly easily by numerical methods. But it would be useful if the condition $\mathcal{C}(\alpha;\,\gamma,\,\theta) = \emptyset$ could be checked in an even simpler way. To do this, we shall derive bounds on the distribution of LR_Q and $\mathcal{W}_Q$. This is done in the two following theorems. In the first one of these, we consider tests for general (nonlinear) restrictions of the form $\mathcal{H}_{NL}$ in (3.9).

Theorem 4.3 A GENERAL BOUND ON THE NULL DISTRIBUTION OF WALD AND LR NONLINEAR TEST STATISTICS. *Under* (2.2), (2.8) *and* (3.9), *the distributions of* LR_{NL} *and* $\mathcal{W}_{NL}$ [*defined in* (3.10) - (3.11)] *may be bounded as follows: for all* $(B,\, J)$,

$$\mathsf{P}_{(B,\,J)}[LR_{NL} \;\; \geq \;\; x] \leq \sup_{(B,\,J)\in H(C_0)} \mathsf{P}_{(B,\,J)}\big[-T\sum_{i=1}^{l}\ln[1-\mu_i(C_0)] \geq x\big]\,,\ \forall x, \quad (4.16)$$

$$\mathsf{P}_{(B,\,J)}[\mathcal{W}_{NL} \;\; \geq \;\; x] \leq \sup_{(B,\,J)\in H(C_0)} \mathsf{P}_{(B,\,J)}\big[T\sum_{i=1}^{l}\frac{\mu_i(C_0)}{1-\mu_i(C_0)} \geq x\big]\,,\ \forall x, \quad (4.17)$$

where $\mathsf{P}_{(B,\,J)}$ *represents the distribution of* Y *when the parameters are* $(B,\, J)$, $H(C_0) = \{(B,\, J) : C_0 B = 0\}$, $C_0 \in \mathcal{M}(k-r,\, k)$, $l = \min\{k-r,\, n\}$, *and* $\mu_1(C_0) \geq \mu_2(C_0) \geq \cdots \geq \mu_n(C_0)$ *are the eigenvalues of the matrix*

$$\bar{F}_L(C_0) = [W'M(C_0)W]^{-1}[W'M(C_0)W - W'MW]\,, \quad (4.18)$$

with M, $M(C_0)$ *and* W *as defined in* (2.13), (3.4) *and* (2.2), *respectively.*

PROOF. Let $H_{NL} = \{(B,\, J) : CB = 0 \text{ for some } C \in \mathcal{M}(k-r,\, k)\}$. It is clear that

$$H_{NL} = \bigcup_{C_0\in\mathcal{M}(k-r,\,k)} H(C_0)\,. \quad (4.19)$$

From Gouriéroux, Monfort and Renault (1993, 1995), we have:

$$LR_{NL} = \inf_{C_0\in\mathcal{M}(k-r,\,k)} LR(C_0), \quad \mathcal{W}_{NL} = \inf_{C_0\in\mathcal{M}(k-r,\,k)} \mathcal{W}(C),$$

which implies that $LR_{NL} \leq LR(C_0)$ and $\mathcal{W}_{NL} \leq \mathcal{W}(C_0)$, for any $C_0 \in \mathcal{M}(k-r,\, k)$. This entails that, for each $C_0 \in \mathcal{M}(k-r,\, k)$ and for all $(B,\, J) \in H(C_0)$,

$$\mathsf{P}_{(B,\,J)}[LR_{NL} \;\; \geq \;\; x] \leq \mathsf{P}_{(B,\,J)}\big[LR(C_0) \geq x\big]\,,\ \forall x, \tag{4.20}$$

$$\mathsf{P}_{(B,\,J)}[\mathcal{W}_{NL} \;\; \geq \;\; x] \leq \mathsf{P}_{(B,\,J)}\big[\mathcal{W}(C_0) \geq x\big]\,,\ \forall x. \tag{4.21}$$

Furthermore, from Theorem **4.1**, we see that, under the null hypothesis $\mathcal{H}_{NL}$, $LR(C_0)$ and $\mathcal{W}(C_0)$ are distributed (respectively) like the variables $-T\sum_{i=1}^{l}\ln(1-\mu_i(C_0))$ and $T\sum_{i=1}^{l}\frac{\mu_i(C_0)}{1-\mu_i(C_0)}$ where $\mu_1(C_0) \geq \mu_2(C_0) \geq \cdots \geq \mu_n(C_0)$ are the eigenvalues underlying (4.18) and $l = \min\{k-r,\, n\}$. From there on, (4.16)-(4.17) follow straightforwardly. □

Theorem **4.3** shows that the distribution of LR_{NL} can be bounded by the distributions of the test statistics $\overline{LR}(C_0)$ [refer to (4.1)-(4.2)] over the set of possible values for C_0. In particular, if the distribution of $\overline{LR}(C_0)$ is the same for all C_0, this yields a single bounding distribution. The latter situation obtains in the Gaussian case (2.3), where an F-distribution can be used. When $k-r=1$, Theorem **4.4** and (4.4) yield the following probability inequality:

$$\mathsf{P}\big[([T-(k-1)-n]/n)\,(LR_{NL}-1) \geq x\big] \leq \mathsf{P}[F(n, T-(k-1)-n) \geq x],\ \forall x. \tag{4.22}$$

Applying these distributional results to the problem at hand leads to the following theorem.

Theorem 4.4 A GENERAL BOUND ON THE NULL DISTRIBUTIONS OF QMM NON-LINEAR TEST STATISTICS. *Under* (2.2), (2.5) *and* (2.7) - (2.9), *the null distribution of* LR_Q *and* $\mathcal{W}_Q$ [*defined by* (3.19)] *may be bounded as follows: for all* $(B,\, J)$

$$\mathsf{P}_{(B,\,J)}[LR_Q \;\; \geq \;\; x] \leq \sup_{(B,\,J)\in H(\gamma_0,\,\theta_0)} \mathsf{P}_{(B,\,J)}\big[-T\ln[1-\mu(\gamma_0,\,\theta_0)] \geq x\big],\ \forall x, \tag{4.23}$$

$$\mathsf{P}_{(B,\,J)}[\mathcal{W}_Q \;\; \geq \;\; x] \leq \sup_{(B,\,J)\in H(\gamma_0,\,\theta_0)} \mathsf{P}_{(B,\,J)}\left[T\frac{\mu(\gamma_0,\,\theta_0)}{1-\mu(\gamma_0,\,\theta_0)} \geq x\right],\ \forall x, \tag{4.24}$$

where $\mathsf{P}_{(B,\,J)}$ *represents the distribution of* Y *when the parameters are* $(B,\, J)$, $H(\gamma_0,\, \theta_0) =$

$\{(B,\ J) : [1,\ \gamma_0,\ \theta_0]\, B = 0\}$*, and* $\mu(\gamma_0,\ \theta_0)$ *is the non-zero eigenvalue of the matrix*

$$\bar{F}_L(\gamma_0,\ \theta_0) = [W'M(\gamma_0,\ \theta_0)W]^{-1}[W'M(\gamma_0,\ \theta_0)W - W'MW]\,, \tag{4.25}$$

where M, $M(\gamma_0,\ \theta_0)$ *and* W *are given by* (2.13), (3.7) *and* (2.2)*, respectively.*

Theorem **4.4** shows that the distribution of LR_Q can be bounded by the distributions of the test statistics $\overline{LR}(\gamma_0,\ \theta_0)$ [refer to (4.5)-(4.6)] over the set of possible values $(\gamma_0,\ \theta_0)$. A single bounding distribution obtains if the distribution of $\overline{LR}(\gamma_0,\ \theta_0)$ is the same for all $(\gamma_0,\ \theta_0)$; in particular, for the Gaussian case, (4.22) yields the following inequality:

$$\mathsf{P}\big[([T-2-n]/n)\,(LR_Q-1) \geq x\big] \leq \mathsf{P}[F(n, T-2-n) \geq x],\ \forall x. \tag{4.26}$$

For non-Gaussian error distributions, the situation is more complex. Again, if the bounding test statistics $\overline{LR}(\gamma_0,\ \theta_0)$ have the same distribution [under $\mathcal{H}_Q(\gamma_0, \theta_0)$] irrespective of $(\gamma_0,\ \theta_0)$, we can obtain a level correct MC bound test by simulating the distribution of $\overline{LR}(\gamma_0,\ \theta_0)$. For generality and further reference, we show here how the generic algorithm associated with a statistic of the general form $S(y,\ X)$ [refer to (4.7)] presented in section 4.1 above can be adapted to yield a bound MC test. So suppose now that $S(y,\ X)$ is not pivotal but is bounded by a pivotal quantity and (4.7) no longer holds, however

$$S(y,\ X) \leq \overline{S}\,(W,\ X)\,.$$

Repeat steps 1-4 above, so the terms $\overline{S}(W^j,\ X)$ are now draws from the bounding distribution; to emphasize this distinction (relative to the pivotal statistics case), we refer to the bound based p-value in step 3 as $\hat{p}_N^{BMC}(S_0)$ or $\hat{p}_N^{BMC}(S_0|\nu)$ when relevant; the corresponding bound critical region from (4.9) is level correct, in the sense that under the null hypothesis $\mathsf{P}\left[\hat{p}_N^{BMC}(S_0) \leq \alpha\right] \leq \alpha$ and $\mathsf{P}\left[\hat{p}_N^{BMC}(S_0|\nu) \leq \alpha\right] \leq \alpha$ for known ν.

This algorithm can be applied to LR_Q using the pivotal expression for $\overline{LR}(\gamma_0,\ \theta_0)$ from (4.5). We denote the associated bound p-value $\hat{p}_N^{BMC}(LR_Q|\gamma, \theta, \nu)$ where: (i) the superscript BMC

[rather than PMC as in (4.10)] indicates that the MC p-value is bound-based, (ii) ν refers to the parameter of the error distribution which is for the moment considered known, and (iii) conditioning on γ and θ identifies the specific choice for γ and θ underlying the bound. This emphasizes that invariance of the bounding distributions generated by $\overline{LR}(\gamma_0,\, \theta_0)$ may not necessarily hold. In this case, numerical methods may be needed to evaluate the bounds in (4.23)-(4.24). So it is not clear that this is any simpler than drawing the set $\mathcal{C}(\alpha;\, \gamma,\, \theta)$ using numerical methods. However, even then, a useful bound may still be obtained by considering the simulated distribution obtained on replacing $(\gamma_0,\, \theta_0)$ by their QMLE estimates in $\overline{LR}(\gamma_0,\, \theta_0)$, where the estimates are treated as fixed known parameters. This yields the following bound p-value:

$$p_N^{BMC}(\hat{\gamma}_Q, \hat{\theta}_Q, \nu) \equiv \hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q, \hat{\theta}_Q, \nu). \tag{4.27}$$

Then, we have the following implication:

$$\hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q, \hat{\theta}_Q, \nu) > \alpha \Rightarrow \hat{p}_N^{PMC}[LR(\gamma_0,\, \theta_0)|\nu] > \alpha\,, \text{ for some } (\gamma_0,\, \theta_0). \tag{4.28}$$

In other words, if $\hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q, \hat{\theta}_Q, \nu) > \alpha$, we can be sure that $\mathcal{C}(\alpha;\, \gamma,\, \theta)$ is *not empty* so that $\mathcal{H}_Q$ is not rejected by the test in (4.14).

4.3. A bootstrap-type LR test

With N draws from the distribution of $W_1,\, \ldots\,,\, W_T$, N simulated samples can be constructed conformably with (2.7)-(2.9), (2.2) under the null hypothesis (2.5), given any value for $\{B, \Sigma\} \in \Psi_Q$ where Ψ_Q refers to the parameter space compatible with the null hypothesis $\mathcal{H}_Q$. Applying (3.19) to each of these samples yields N simulated statistics from the null data generating process. Count the number of simulated statistics $\geq$ {observed LR_Q} and replace this number for $\hat{G}_N(.)$ in (4.8); this provides a MC p-value conditional on the choice for B and Σ which we denote $\hat{p}_N^{LMC}(LR_Q|B, \Sigma, \nu)$ where the symbol LMC stands for *Local* MC to emphasize that it is estimated given a specific nuisance parameter value. In particular, it is natural to consider the QMLE estimates $\hat{B}_Q$ and $\hat{\Sigma}_Q$ defined in (3.20) which yields

$$p_N^{LMC}(\hat{B}_Q, \hat{\Sigma}_Q, \nu) \equiv \hat{p}_N^{LMC}(LR_Q|\hat{B}_Q, \hat{\Sigma}_Q, \nu). \tag{4.29}$$

This bootstrap-type p-value differs from the bound-based p-value $\hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q, \hat{\theta}_Q, \nu)$ from (4.27), although both use QMLE estimates. Indeed, as a consequence of Theorem **4.4**, it is straightforward to see that $\hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q, \hat{\theta}_Q, \nu) \geq \hat{p}_N^{LMC}(LR_Q|\hat{B}_Q, \hat{\Sigma}_Q, \nu)$. Relying on the latter will not necessarily yield a level correct test in finite samples. However, since

$$\hat{p}_N^{LMC}(LR_Q|\hat{B}_Q, \hat{\Sigma}_Q, \nu) > \alpha \Rightarrow \sup_{\{B,\Sigma\}\in\Psi_Q} \hat{p}_N^{LMC}(LR_{BCAPM}|B, \Sigma, \nu) > \alpha$$

where B, Σ refer to the parameter space compatible with the null hypothesis. So a non-rejection based on $\hat{p}_N^{LMC}(LR_Q|\hat{B}_Q, \hat{\Sigma}_Q, \nu)$ can be considered conclusive from an exact test perspective.

5. The case of error distributions with unknown parameters

We will now extend the above results to the unknown distributional parameter case for the error families of particular interest, namely (2.2). To do this, it is helpful to first revisit the generic algorithm and its modification presented in sections 4.1 and 4.2 above. The outcome of the both algorithms, namely the p-values $\hat{p}_N^{PMC}(S_0|\nu)$ and $\hat{p}_N^{BMC}(S_0|\nu)$ are valid if ν is known. When ν is unknown, we can apply the method of maximized Monte Carlo (MMC) tests [see Dufour (2006)] to obtain tests that satisfy the level constraint even in finite samples: given a set of nuisance parameters Φ_0 consistent with the null hypothesis of interest, the critical regions

$$\sup_{\nu\,\in\,\Phi_0} [\hat{p}_N^{PMC}(S_0|\nu)] \leq \alpha, \qquad \sup_{\nu\,\in\,\Phi_0} [\hat{p}_N^{BMC}(S_0|\nu)] \leq \alpha. \tag{5.1}$$

have level α. This suggests to maximize the MC p-values presented in sections 4.1, 4.2 and 4.3 above. Rather than considering Φ_0 as a search set, we focus on a set estimate for ν [see e.g. Dufour and Kiviet (1996) and Beaulieu, Dufour and Khalaf (2005, 2006, 2007)] to ensure that the error distributions retained are empirically relevant. Formally, we consider the following critical regions:

$$Q_U^{BMC}(\gamma,\theta) \leq \alpha_2\,, \quad Q_U^{BMC}(\hat{\gamma}_Q,\hat{\theta}_Q) \leq \alpha_2\,, \tag{5.2}$$

$$Q_U^{BMC}(\hat{\gamma}_Q,\hat{\theta}_Q) = \sup_{\nu\,\in\,\mathcal{C}(Y)} \hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q,\hat{\theta}_Q,\nu), \tag{5.3}$$

$$Q_U^{LMC}(\hat{B}_Q,\hat{\Sigma}_Q) = \sup_{\nu\in\,\mathcal{C}(Y)} \hat{p}_N^{LMC}(LR_Q|\hat{B}_Q,\hat{\Sigma}_Q,\nu) \tag{5.4}$$

where $\hat{p}_N^{BMC}(LR_Q|\hat{\gamma}_Q,\hat{\theta}_Q,\nu)$ [defined in (4.27)] is the specific bound based on QMLE estimates of γ and θ, $\hat{p}_N^{LMC}(LR_Q|\hat{B}_Q,\hat{\Sigma}_Q,\nu)$ [defined in (4.29)] is the LMC bootstrap-type p-value, and $\mathcal{C}(Y)$ is an $(1\text{-}\alpha_1)$ level confidence set for ν [described below]. Non-rejections [refer to sections 4.2 and 4.3] based on (5.2) are exact at the $\alpha_1+\alpha_2$ level. In section 8, we use $\alpha_1=\alpha_2=\alpha/2$.

To obtain $\mathcal{C}(Y)$, we "invert" a test for hypothesis (2.2) where $\nu=\nu_0$ and known ν_0. Inverting an α_1 level test involves collecting the values of ν_0 not rejected by this test at level α_1. We consider a test which assesses lack-of-fit of the hypothesized distribution. In what follows, we briefly describe this test; for more complete algorithms, proofs and references, see Dufour, Khalaf and Beaulieu (2003) and Beaulieu et al. (2007). We consider

$$\mathrm{ESK}(\nu_0) = \left|\mathrm{SK} - \overline{\mathrm{SK}}(\nu_0)\right|, \qquad \mathrm{EKU}(\nu_0) = \left|\mathrm{KU} - \overline{\mathrm{KU}}(\nu_0)\right|, \tag{5.5}$$

where SK and KU are the multivariate skewness and kurtosis criteria

$$\mathrm{SK} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{i=1}^{T}\hat{d}_{ii}^3, \qquad \mathrm{KU} = \frac{1}{T}\sum_{t=1}^{T}\hat{d}_{tt}^2, \tag{5.6}$$

$\hat{d}_{it}$ are the elements of the matrix $\hat{U}(\hat{U}'\hat{U})^{-1}\hat{U}'$, and $\overline{\mathrm{SK}}(\nu_0)$ and $\overline{\mathrm{KU}}(\nu_0)$ are simulation-based estimates of the expected SK and KU given (2.2). These are obtained, given ν_0, by drawing samples conformable with (2.2) then computing the corresponding average measures of skewness and kurtosis. The MC test technique may be applied to obtain exact p-values for $\mathrm{ESK}(\nu_0)$ and $\mathrm{EKU}(\nu_0)$ as follows. Conditional on the same $\overline{\mathrm{SK}}(\nu_0)$ and $\overline{\mathrm{KU}}(\nu_0)$, generate, imposing $\nu=\nu_0$, replications of the excess skewness and excess kurtosis statistics. Exact Monte Carlo p-values [denoted

$\hat{p}(\mathrm{ESK}_0\,|\nu_0)$ and $\hat{p}(\mathrm{EKU}_0\,|\nu_0)$] for each test statistic can be calculated from the rank of the observed $\mathrm{ESK}(\nu_0)$ and $\mathrm{EKU}(\nu_0)$ relative to the simulated ones. The generic algorithm presented in section 4.1 also applies here, where the relevant pivotal characterizations can be found in Dufour et al. (2003). To obtain a joint test based on SK and KU, we consider the joint criterion:

$$\mathrm{CSK} = 1 - \min\left\{\hat{p}(\mathrm{ESK}_0\,|\nu_0), \hat{p}(\mathrm{EKU}_0\,|\nu_0)\right\}. \tag{5.7}$$

The MC test technique is once again applied to obtain a size correct p-value for the combined test.

6. A less restricted model

We now turn to the general hypothesis (2.14). Let us first observe that the hypothesis may be rewritten in matrix form $\mathcal{H}_{GQMM} : [1,\,\gamma,\,\theta]\,B = \phi\iota_n'$, where ι_n is an n-dimensional vector of ones. If ϕ is known, then it is possible to use the above eigenvalue test procedure to derive its associated LR statistic. One may *e.g.* consider the MLR:

$$R_{it} - \tilde{R}_{\mathrm{M}t} - \phi = \overline{a}_i + (\overline{b}_i - 1)\tilde{R}_{\mathrm{M}t} + \overline{c}_i\tilde{R}_{\mathrm{M}t}^2 + u_{it},\ t = 1,\,\ldots\,,T,\ i = 1,\,\ldots\,,n. \tag{6.8}$$

In this context, testing $\mathcal{H}_{GQMM}$ corresponds to assessing

$$\overline{a}_i + \gamma(\overline{b}_i - 1) + \overline{c}_i\theta = 0, \quad i = 1,\,\ldots\,,n, \tag{6.9}$$

which leads to the above framework. This also suggests a simple procedure to derive the LR statistic to test $\mathcal{H}_{GQMM}$ and the associated MLE estimate of ϕ. Indeed, one may minimize, over ϕ, the eigenvalue based criterion associated with (6.8)-(6.9); this may be conducted numerically (yet since the argument of the underlying minimization is scalar, this approach is numerically much more tractable than the iterative maximization applied by Barone-Adesi et al. (2004*b*)). The statistical solutions we derived above also hold in this framework: indeed, bounding cut-off points can be obtained using the null distribution of the test statistic which fixes ϕ to a known value. Note however that we have shown in Dufour and Khalaf (2002) that the latter distribution does not depend on ϕ, which will lead one to the same bound whether ϕ is zero or not. Tighter bounds can be obtained under normality; see Velu and Zhou (1999) or Zhou (1995).

7. Simulation study

We now present a small-scale simulation experiment to assess the performance of the proposed tests. We focus on $\mathcal{H}_Q$ [hypothesis (2.5)] in the context of (2.1). The design is calibrated to match our empirical analysis (see section 8) which relies on monthly returns of 25 value-weighted portfolios from Fama and French's data base, for 1961-2000. Since we assess the model over 5-year intervals (as well as over the whole sample), we consider (2.1) with $n = 25$ equations and $T = 60$, which reflect our subperiod analysis. $\tilde{R}_{\mathrm{M}t}$, $t = 1, ..., T$ are fixed to the returns on the market portfolio from the aforementioned data set for the 1996-2000 time period. The coefficients of this regressor and its square, namely b_i and c_i, $i = 1, \ldots, n$, in (2.1), are fixed to their observed counterparts, namely, $\hat{b}_i$ and $\hat{c}_i$, $i = 1, \ldots, n$, corresponding to the unconstrained OLS regression over the 1996-2000 sample period; from this same regression, we also retain: (i) the largest intercept estimate

$$\overline{\overline{a}} = \max_i \{\hat{a}_i\}, \quad i = 1, \ldots, n, \tag{7.1}$$

to set the scale for the power study (see below), and (ii) the estimated variance-covariance error matrix, to generate model shocks; formally, we use (2.2) substituting the Cholesky factor of the variance-covariance estimate in question for the J matrix.

We study normal and t-errors with unknown degrees-of-freedom, so the random errors W_t, $t = 1, ..., T$ in (2.2) are generated as in (2.3) and (2.4), respectively. In the latter case, the degrees-of-freedom parameter κ is fixed to 8. This choice is also motivated by our empirical application: as may be checked from Table 4 in section 8, the average lower limit of the confidence set for κ over the eight 5-year subperiods analyzed is around 8. The MC tests are applied imposing and ignoring information on κ, which allows to document the cost of estimating this parameter. When κ is taken as unknown, MMC p-values are calculated over the space $4 \leq \kappa \leq 13$. A wider range is allowed in our empirical application; in the case of the simulation study, this restriction is adopted to keep execution time manageable.

Several choices for the key parameters γ and θ are evaluated, particularly to assess the size of asymptotic tests. Again, we rely on the data set analyzed in the next section, as follows. We consider, in turn, the QMLE estimates of γ and θ from each of the eight 5-year sub-samples analyzed; we also consider an alternative design setting γ and θ to the average of their estimates over the eight 5-year subperiods. The latter design is used for the power study (see below).

We study the LR statistic defined in the previous sections, with its asymptotic χ^2 p-value, its QMLE-based bound p-value and its QMLE-based bootstrap type p-value.

The model intercepts are set as follows. For the size study, $\mathcal{H}_Q$ [hypothesis (2.5)] is imposed so a_i, $i = 1, \ldots, n$, are obtained given the choices for b_i and c_i, and γ and θ described above. For the power study, intercepts are generated such that

$$a_i = -\gamma(b_i - 1) - c_i\theta + \Delta\,\overline{\overline{a}}, \quad i = 1, \ldots, n, \tag{7.2}$$

where $\overline{\overline{a}}$ is defined in (7.1) and Δ is a scalar ranging from 0.25 to 2.5 which controls the extent of departure from $\mathcal{H}_Q$; clearly, $\Delta = 0$ yields $\mathcal{H}_Q$. Here again, the constant $\overline{\overline{a}}$ aims at calibrating the design to our empirical application, for empirical relevance purposes.

In all designs, $N = 99$ replications are used to implement MC tests [we used 999 in the empirical application]. The literature on MC testing [see *e.g.* Dufour and Kiviet (1996), Dufour and Khalaf (2002) or Dufour, Khalaf, Bernard and Genest (2004)] illustrates reliability with $N = 99$ in simulation studies. In each experiment, the number of simulations is 1000; we report empirical rejections for a nominal level of 5%.

Results reveal that the χ^2 asymptotic test is oversized for all designs considered; indeed, empirical sizes range from a minimum of 25.8% to 32.3%. Since we have attempted to calibrate simulation designs to the empirical study, these results confirm our reliance on MC alternatives to the usual asymptotic cut-off points. It is worth noting that over-rejections with normal errors are almost as severe as with the Student-t case. To interpret the bootstrap results in Table 1, recall that for the Student-t distribution, the degrees-of-freedom are assumed known (this assumption is relaxed in Table 2), whereas all remaining nuisance parameters are estimated consistently under the

Table 1. Size of asymptotic and parametric bootstrap QMM tests

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Normal	Asy	.287	.291	.279	.295	296	.275	.258	.287
	LMC	.062	.062	.051	.087	.072	.039	.029	.077
Student-t	Asy	.314	.313	.317	.300	.314	.323	.288	.319
	LMC [κ known]	.073	.074	.069	.087	.077	.062	.029	.109

Note – Numbers shown are empirical rejections for proposed tests of $\mathcal{H}_Q$ [hypothesis (2.5) in the context of (2.1)]. The statistic considered is the quasi-LR statistic (3.19); associated p-values rely on, respectively, the asymptotic $\chi^2(n-2)$ distribution and the LMC p-values assuming the error distribution is known, which corresponds to a parametric bootstrap. Columns (1) - (8) refer to the choices for the parameters γ and θ underlying the various simulation designs considered. These parameters are fixed, in turn, to their QMLE counterparts based on the data set analyzed in section 8, over each of the eight 5-year subperiods under study, so column (1) refers to parameters estimated using the 1961-65 subsample, column (2) to the 1966-70 subsample, etc.

null hypothesis. Bootstrapping reduces over-rejections so tests based on LMC p-values deviate only moderately from their nominal size: rejection probabilities range from 2.9% to 8.7% with normal errors and from 2.9% to 10.9% with Student-t errors.

Results of the power study reported in Table 2 show that the tests have a good power performance. Observe that empirical rejections associated with $\Delta \neq 0$ in columns (1) and (4) of Table 2 convey a misleading assessment of power, since the underlying χ^2-based tests are severely oversized. In such cases, a size-correction scheme is required; for instance, one may compute an artificial size-correct cut-off point from the quantiles of the simulated statistics conditional on each design. Since the bootstrap-type correction seems to work, at least locally, for the design under study, we prefer to analyze the bootstrap version of the tests rather than resorting to another artificial size correction. Indeed, whereas bootstrap-based size corrections are empirically applicable, a local correction corresponds to a practically infeasible test. Though simulation results generally depend on the designs considered, the following findings summarized next are worth noting.

1. Estimation costs for the degrees-of-freedom parameter with Student-t errors are unnoticeable. Indeed, the empirical rejections based on BMC p-values are identical whether κ is treated as a known or as an unknown scalar. The bootstrap type LMC tests are affected albeit moderately from estimating κ.

Table 2. Size and power of QMM tests

	Normal				Student-t			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ	Asy	BMC	LMC	asy	κ known		κ unknown	
					BMC	LMC	BMC	LMC
0	.288	.025	.052	.320	.044	.069	.044	.065
.25	.571	.109	.184	.533	.097	.140	.097	.134
.50	.915	.444	.562	.840	.261	.385	.261	.377
.75	.977	.718	.812	.949	.418	.606	.418	.591
1.0	.987	.807	.881	.971	.508	.711	.508	.701
1.5	.995	.848	.916	.981	.581	.795	.581	.784
2.0	.997	.862	.925	.983	.608	.808	.608	.796
2.5	.997	.866	.924	.983	.618	.821	.618	.811

Note – Numbers shown are empirical rejections for proposed tests of $\mathcal{H}_Q$ [hypothesis (2.5) in the context of (2.1)]. The null hypothesis corresponds to $\Delta = 0$. The statistic considered is the quasi-LR statistics [refer to (3.19)]; associated p-values rely on, respectively, the asymptotic $\chi^2(n-2)$ distribution [columns (1) and (4)], the BMC and LMC p-values assuming the error distribution is known [columns (2,3) and (5,6)], and the BMC and LMC p-values imposing multivariate $t(\kappa)$ errors with unknown κ [columns (7,8)]; the latter p-values are the largest over the degrees-of-freedom parameter κ within the specified search set.

2. Size-correct tests seem to perform well for the considered alternative [see (7.2)], which focuses on homogenous deviations from $\mathcal{H}_Q$. Such alternatives may be harder to detect relative to the nonhomogeneous case; our findings thus provide a realistic appraisal of power given a possibly less favorable (though empirically relevant) scenario.

3. Tests based on LMC p-values outperform bound tests based on BMC cutoffs. We observe power differences averaging around 20% with normal errors, 38% with Student-t errors and known degrees-of-freedom, and 37% with Student-t errors and unknown degrees-of-freedom. The BMC test is however not utterly conservative. Taken collectively, size and power rankings emerging from this study illustrate the reliability of our proposed test strategy which relies on the BMC in conjunction with the LMC p-values.

4. Controlling for nuisance parameter effects, it seems that kurtosis in the data reduces power. Two issues need to be raised in this regard. First, recall that both LMC and BMC p-values are obtained using the Gaussian QMLE estimates of γ and θ and both LMC and

BMC procedures are not invariant to the latter. In addition, the test statistic is based on Gaussian-QMLE, although we have corrected its critical region for departures from normality. Gaussian-QMLE coincides with least squares in this model, and least-squares-based statistics are valid (although possibly not optimal) in non-normal settings, at least in principle. We view our results as a motivation for research on finite-sample robust test procedures in MLR models.

8. Empirical analysis

For our empirical analysis, we use Fama and French's data base. We produce results for monthly returns of 25 value-weighted portfolios from 1961-2000. The portfolios which are constructed at the end of June, are the intersections of five portfolios formed on size (market equity) and five portfolios formed on the ratio of book equity to market equity. The size breakpoints for year s are the New York Stock Exchange (NYSE) market equity quintiles at the end of June of year s. The ratio of book equity to market equity for June of year s is the book equity for the last fiscal year end in $s-1$ divided by market equity for December of year $s-1$. The ratio of book equity to market equity are NYSE quintiles. The portfolios for July of year s to June of year $s+1$ include all NYSE, AMEX, NASDAQ stocks for which we have market equity data for December of year $s-1$ and June of year s, and (positive) book equity data for $s-1$. All MC tests where applied with 999 replications, and multivariate normal and multivariate Student-t errors; formally, as in (2.3) and (2.4) respectively.

Table 3 reports tests of $\mathcal{H}_Q$ [hypothesis (2.5) in the context of (2.1)] and of $\mathcal{H}_{GQMM}$ [hypothesis (2.14) in the context of (2.1)], over intervals of 5 years and over the whole sample. Subperiod analysis is usual in this literature [see *e.g.* the surveys of Black (1993) or Fama and French (2004)], and is mainly motivated by structural stability arguments. Our previous and ongoing work on related asset pricing applications [Beaulieu, Dufour and Khalaf (2005, 2006, 2007), Dufour, Khalaf and Beaulieu (2008)] have revealed significant temporal instabilities which support subperiod analysis even in conditional models which allow for time varying *betas*.

To validate our statistical setting, companion diagnostic tests are run and are reported in Table 4. These include: (i) goodness of fit tests associated with distributional hypotheses (2.3) and (2.4);

Table 3. QMM and GQMM tests

	Tests of $\mathcal{H}_Q$					Tests of $\mathcal{H}_{GQMM}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	Asy	$\hat{p}_{\mathcal{N}}^{BMC}$	$\hat{p}_{\mathcal{N}}^{LMC}$	$\hat{p}_{\mathsf{t}}^{BMC}$	$\hat{p}_{\mathsf{t}}^{LMC}$	Asy	$\hat{p}_{\mathcal{N}}^{BMC}$	$\hat{p}_{\mathcal{N}}^{LMC}$	$\hat{p}_{\mathsf{t}}^{BMC}$	$\hat{p}_{\mathsf{t}}^{LMC}$
1961-1965	.078	.586	.405	.596	.414	.099	.666	.468	.678	.474
1966-1970	.008	.246	.134	.255	.143	.217	.813	.646	.828	.652
1971-1975	.054	.515	.341	.529	.367	.227	.815	.664	.828	.674
1976-1980	.000	.006	.002	.015	.006	.215	.793	.670	.810	.679
1981-1985	.001	.081	.022	.093	.033	.002	.133	.061	.176	.085
1986-1990	.003	.163	.074	.193	.091	.029	.448	.283	.467	.295
1991-1995	.001	.114	.055	.148	.076	.002	.163	.071	.196	.096
1996-2000	.073	.543	.386	.591	.420	.229	.803	.665	.830	.677
1961-2000	.000	.001	.001	.002	.002	.003	.011	.015	.002	.005

Note – Columns (1)-(5) pertain to tests of $\mathcal{H}_Q$ [hypothesis (2.5) in the context of (2.1)]; columns (6)-(10) pertain to tests of $\mathcal{H}_{GQMM}$ [hypothesis (2.14) in the context of (2.1)]. Numbers shown are *p*-values, associated with the quasi-LR statistics [refer to (3.19)], relying on, respectively, the asymptotic $\chi^2(n-2)$ distribution [columns (1) and (6)], the Gaussian based BMC and LMC p-values [columns (2)-(3) and (7)-(8)], and the BMC and LMC p-values imposing multivariate $t(\kappa)$ errors [columns (4)-(5) and (9)-(10)]. MC p-values for $t(\kappa)$ errors are the largest over the degrees-of-freedom parameter κ within the specified confidence sets; the latter is reported in column 6 of Table 4. January and October 1987 returns are excluded from the dataset.

(ii) tests for departure from the maintained error *i.i.d.* hypothesis, and (iii) tests for exogeneity of the market factors.

The goodness-of-fit tests rely on the multivariate skewness-kurtosis criteria described in section 5. For the normal distribution, we apply the pivotal MC procedure to the omnibus statistic:

$$MN = \frac{T}{6}\,\mathrm{SK} + \frac{T\,[\mathrm{KU} - n(n+2)]^2}{8n(n+2)}\,. \tag{8.1}$$

For the Student-t distribution, we report the confidence set for the degrees-of-freedom parameter which inverts the combined skewness-kurtosis statistic (5.7).

Serial dependence tests [from Dufour et al. (2008) and Beaulieu et al. (2007)] are summarized here for convenience. In particular, we apply the LM-GARCH test statistic [Engle (1982)] and the variance ratio statistic which assesses linear serial dependence [Lo and MacKinlay (1988)], to standardized residuals, namely $\tilde{W}_{it}$, the elements of the matrix

$$\tilde{W} = \hat{U}\, S_{\hat{U}}^{-1}, \tag{8.2}$$

where $S_{\hat{U}}$ is the Cholesky factor of $\hat{U}'\hat{U}$. So the modified GARCH test statistic for equation i, denoted E_i, is given by $T\times$ (the coefficient of determination in the regression of the equation's squared OLS residuals $\tilde{W}_{it}^2$ on a constant and $\tilde{W}_{(t-j),i}^2$, $j = 1, \ldots , q$) where q is the ARCH order against which the test is designed. The modified variance ratio is given by:

$$\tilde{VR}_i = 1 + 2\sum_{j=1}^{K}\Big(1 - \frac{j}{K}\Big)\hat{\rho}_{ij}\,, \quad \hat{\rho}_{ij} = \frac{\sum_{t=j+1}^{T}\tilde{W}_{it}\tilde{W}_{i,t-j}}{\sum_{t=1}^{T}\tilde{W}_{ti}^2}\,. \tag{8.3}$$

12 lags are used for both procedures. We combine inference across equation via the joint statistics:

$$\tilde{E} = 1 - \min_{1\le i\le n}\big[p(\tilde{E}_i)\big]\,, \qquad \tilde{VR} = 1 - \min_{1\le i\le n}\big[p(\tilde{VR}_i)\big]\,, \tag{8.4}$$

where $p(\tilde{E}_i)$ and $p(\tilde{VR}_i)$ refer to p-values, obtained using the $\chi^2(q)$ and $N\big[1,\, 2(2K-1)(K-1)/(3K)\big]$ respectively. In Dufour et al. (2008), we show that under (2.2), $\tilde{W}$ has a distribution which depends only on κ, so the MC test technique can be applied to obtain a size correct p-value for $\tilde{E}$ and $\tilde{VR}$. To deal with an unknown κ, we apply an MMC test procedure following the same technique proposed for tests on $\mathcal{H}_Q$. Specifically, we use the same confidence set for κ, of level $(1-\alpha_1)$; we maximize the p-value function associated with $\tilde{E}$ and $\tilde{VR}$ over all values of κ in the latter confidence set; we then refer the latter maximal p-value to α_2 where $\alpha = \alpha_1 + \alpha_2$. Power properties of these tests are analyzed in Dufour et al. (2008) and suggest a good performance for sample sizes compatible with our subperiod analysis.

We also apply the Wu-Hausman test to assess the potential endogeneity of our regressors. It consists in appending, to each equation, the residuals from a first stage regression of returns on a constant and the instruments, and testing for the exclusion of these residuals using the usual OLS based F-statistic [see Hausman (1978), Dufour (1987)]. This test is run, in turn, for each equation, with one lag of $\tilde{R}_{\mathrm{M}}$, $\tilde{R}_{\mathrm{M}}^2$ and R_i, $i = 1, 25$ as instruments. Numbers shown are the minimum

p-values over all equations. The usual F-based p-value is computed for the normal case; for the Student-t, we compute MMC p-values, as follows. In each equation, and ignoring contemporaneous correlation of the error term, the *F*-statistic in question is location-scale invariant and can easily be simulated to derive a MC p-value given draws from a Student-t distribution, conditional on its degrees-of-freedom. We maximize the p-value so obtained over κ in the same confidence set used for all other tests as described above; we then refer the latter maximal p-value to α_2 where $\alpha = \alpha_1 + \alpha_2$. For presentation clarity, we report the minimum p-value in each case, over all equations.

For all confidence set based MMC tests under the Student-t hypothesis, we consider $\alpha_1 = 2.5\%$ so, in interpreting the p-values reported in following tables for the Student-t case, α_1 must be subtracted from the adopted significance level; for instance, to obtain a 5% test, reported p-values should be referred to 2.5% as a cut-off.

From Table 3, we see that, when assessed using the whole sample, both $\mathcal{H}_Q$ and $\mathcal{H}_{GQMM}$ are soundly rejected, using asymptotic or MC p-values, the confidence sets on the degrees-of-freedom parameter is quite tight and suggests high kurtosis, and normality is definitely rejected. Unfortunately, the diagnostic tests (Table 4) reveal significant departures from the statistical foundations underlying the latter tests (even when allowing for non-normal errors); temporal instabilities thus cast doubt on the full sample analysis.

Results over subperiods can be summarized as follows. Multivariate normality is rejected in many subperiods and provides us with a reason to investigate whether test results shown under multivariate normality are still prevalent once we use Student-t distributions. $\mathcal{H}_Q$ is rejected at the 5% level in five subperiods out of eight using asymptotic p-values. Using finite-sample tests under multivariate normality reveals that $\mathcal{H}_Q$ is rejected at the 5% level in only one subperiod, namely 1976-1980. The LMC p-value confirms all these non-rejections but one. Using the same approach under the multivariate Student-t distribution leads to the same conclusion. $\mathcal{H}_{GQMM}$ is not rejected in any subperiod allowing for *t*-errors, although the normal LMC p-value is less than 5% for 1986-90, and asymptotic p-values are highly significant for three subperiods spanning 1981-1995. Diagnostic tests allowing for t errors reveal significant (at the 5% level) departures from the *i.i.d.*

Table 4. Multivariate diagnostics

	Time Dependence				Goodness-of-fit		Exogeneity	
	$\tilde{E}$		$\tilde{VR}$		MN	$CS(\kappa)$	Wu-Hausman	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Normal	Student-t	Normal	Student-t	Normal	Student-t	Normal	Student-t
1961-1965	.718	.769	.117	.135	.026	≥ 8	.214	.235
1966-1970	.258	.314	.954	.996	.044	≥ 8	.079	.100
1971-1975	.215	.266	.253	.260	.015	≥ 7	.018	.026
1976-1980	.001	.004	.502	.516	.004	≥ 6	.003	.005
1981-1985	.222	.237	.131	.148	.001	≥ 5	.143	.161
1986-1990	.544	.559	.056	.070	.401	≥ 12	.042	.053
1991-1995	.150	.166	.142	.149	.176	≥ 9	.259	.285
1996-2000	.010	.049	.847	.849	.001	$3-11$	.194	.189
1961-2000	.001	.001	.002	.007	.001	$5-8$	.000	.001

Note – Numbers shown in columns (1)-(5) and (7)-(10) are p-values associated with the combined test statistics $\tilde{E}$ [columns (1) and (2)], $\tilde{VR}$ [columns (3) and (4)], MN [column (5)] and $\tilde{WH}$ [columns (7)-(10)]. $\tilde{E}$, defined by (8.4), is a multivariate extension of Engle's GARCH test statistic. $\tilde{VR}$, defined by (8.4), is a multivariate extension of Lo and MacKinlay's variance ratio test. MN is a MC version of the multivariate combined skewness and kurtosis test based on (8.1). The Wu-Hausman test is applied with one lag of $\tilde{R}_{\text{M}}$, $\tilde{R}_{\text{M}}^2$ and R_i, $i = 1, 25$ as instruments. Numbers shown are the minimum p-values over all equations; the usual F-based p-value is computed for the normal case; for the Student-t, we compute MMC p-values. Both normal and Student-t p-values for this test are univariate. In columns (1), (3), (5) and (7), the Gaussian p-values are MC pivotal statistics based; p-values in columns (2), (4), (6) and (8) are MMC confidence set based; the relevant 2.5% confidence set for the nuisance parameters is reported in column (6). Specifically, $CS(\kappa)$ corresponds to the confidence set estimate of level 97.5% for the degrees-of-freedom parameter of the multivariate Student-t error distribution; this set is obtained by inverting the goodness-of-fit statistic (5.7).

hypothesis in the 1976-1980 subperiod and not elsewhere. Recall that in this same subperiod, our QLR tests reject $\mathcal{H}_Q$ at the 5% level.

We conclude by underlying the evidence that contrary to the asset pricing evidence in the literature, this version of the CAPM is generally not rejected by our tests, even when controlling for finite-sample inference. Compared to the results of Beaulieu, Dufour and Khalaf (2005, 2006, 2007), we see that the QMM model is not rejected using our tests, whereas both Black's version of the CAPM, or the CAPM with observed risk-free rate are rejected using related test methods. This observation must be qualified since the diagnostic tests applied to the overall sample are significant at conventional levels revealing temporal instabilities. Care must be exercised in interpreting our generalized Wu-Hausman test results. Indeed, recall that the reported p-values are the smallest over all equations, and the p-values (including the simulated ones), are univariate, in the sense that contemporaneous correlation of shocks is ignored. If we consider a Bonferroni approach to obtain a valid joint test, which involves dividing the adopted test level by 25 here, we see that for the full sample, the test remains significant even with such a conservative correction. This result may be viewed as a motivation for conditional three-moment based modeling.

9. Conclusion

In this paper, we consider the quadratic market model which extends the standard CAPM framework to incorporate the effect of asymmetry of return distribution on asset valuation. The development of exact tests of the QMM is an appealing research objective, given: (i) the increasing popularity of this model in finance, (ii) the fact that traditional market models (which suppose that asset returns move proportionally to the market) have not fared well in empirical tests, (iii) available related studies are only asymptotic (exact tests are unavailable even with normal errors). We have proposed exacts tests of the QMM allowing for non-normal distributions exactly. The underlying statistical challenges relate to dimensionality tests which are interesting in their own right. Our results show that although asymptotic tests are significant in several subperiods, exact tests fail to reject this model with Fama-French data. Temporal instabilities are however evident, and motivate exploring conditional three-moment based models.

References

Barone-Adesi, G. (1985), 'Arbitrage equilibrium with skewed asset returns', *Journal of Financial and Quantitative Analysis* **20**(3), 299–313.

Barone-Adesi, G., Gagliardini, P. and Urga, G. (2004*a*), A test of the homogeneity hypothesis on asset pricing models, *in* E. Jurczenko and B. Maillett, eds, 'Multi-Moment CAPM and Related Topics', Springer-Verlag, Berlin.

Barone-Adesi, G., Gagliardini, P. and Urga, G. (2004*b*), 'Testing asset pricing models with coskewness', *Journal of Business and Economic Statistics* **22**(4), 474–485.

Beaulieu, M.-C., Dufour, J.-M. and Khalaf, L. (2005), Exact multivariate tests of asset pricing models with stable asymmetric distributions, *in* M. Breton and H. Ben Ameur, eds, 'Numerical Methods in Finance', Kluwer/Springer-Verlag, New York, chapter 9, pp. 173–191.

Beaulieu, M.-C., Dufour, J.-M. and Khalaf, L. (2006), Testing Black's CAPM with possibly non-gaussian errors: An exact identification-robust simulation-based approach, Technical report, CIREQ, Université de Montréal and Université Laval.

Beaulieu, M.-C., Dufour, J.-M. and Khalaf, L. (2007), 'Multivariate tests of mean-variance efficiency with possibly non-Gaussian errors: An exact simulation-based approach', *Journal of Business and Economic Statistics* **25**(4), 398–410.

Black, F. (1972), 'Capital market equilibrium with restricted borrowing', *Journal of Business* **45**, 444–454.

Black, F. (1993), 'Beta and return', *Journal of Portfolio Management* **20**(1), 8–17.

Bolduc, D., Khalaf, L. and Moyneur, E. (2008), 'Identification-robust simulation-based inference in joint discrete/continuous models for energy markets', *Computational Statistics and Data Analysis* **forthcoming**.

Campbell, J. Y. (2003), 'Asset pricing at the millennium', *Journal of Finance* **55**, 1515–1567.

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

Dittmar, R. (2002), 'Nonlinear pricing kernels, kurtosis preferences and the cross-section of equity returns', *Journal of Finance* **57**, 369–403.

Dufour, J.-M. (1987), Linear Wald methods for inference on covariances and weak exogeneity tests in structural equations, *in* I. B. MacNeill and G. J. Umphrey, eds, 'Advances in the Statistical Sciences: Festschrift in Honour of Professor V.M. Joshi's 70th Birthday. Volume III, Time Series and Econometric Modelling', D. Reidel, Dordrecht, The Netherlands, pp. 317–338.

Dufour, J.-M. (1997), 'Some impossibility theorems in econometrics, with applications to structural and dynamic models', *Econometrica* **65**, 1365–1389.

Dufour, J.-M. (2003), 'Identification, weak instruments and statistical inference in econometrics', *Canadian Journal of Economics* **36**(4), 767–808.

Dufour, J.-M. (2006), 'Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics in econometrics', *Journal of Econometrics* **133**(2), 443–477.

Dufour, J.-M. and Khalaf, L. (2002), 'Simulation based finite and large sample tests in multivariate regressions', *Journal of Econometrics* **111**(2), 303–322.

Dufour, J.-M., Khalaf, L. and Beaulieu, M.-C. (2003), 'Exact skewness-kurtosis tests for multivariate normality and goodness-of-fit in multivariate regressions with application to asset pricing models', *Oxford Bulletin of Economics and Statistics* **65**, 891–906.

Dufour, J.-M., Khalaf, L. and Beaulieu, M.-C. (2008), 'Multivariate residual-based finite-sample tests for serial dependence and GARCH with applications to asset pricing models', *Journal of Applied Econometrics* **forthcoming**.

Dufour, J.-M., Khalaf, L., Bernard, J.-T. and Genest, I. (2004), 'Simulation-based finite-sample tests for heteroskedasticity and ARCH effects', *Journal of Econometrics* **122**(2), 317–347.

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

Dufour, J.-M. and Taamouti, M. (2005), ‘Projection-based statistical inference in linear structural models with possibly weak instruments’, *Econometrica* **73**(4), 1351–1365.

Dufour, J.-M. and Taamouti, M. (2007), ‘Further results on projection-based inference in IV regressions with weak, collinear or missing instruments’, *Journal of Econometrics* **139**(1), 133–153.

Engle, R. F. (1982), ‘Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation’, *Econometrica* **50**(4), 987–1008.

Fama, E. F. and French, K. R. (2004), ‘The Capital Asset Pricing Model: Theory and evidence’, *The Journal of Economic Perspectives* **18**(3), 25–46.

Gibbons, M. R. (1982), ‘Multivariate tests of financial models: A new approach’, *Journal of Financial Economics* **10**, 3–27.

Gouriéroux, C., Monfort, A. and Renault, E. (1993), ‘Tests sur le noyau, l’image et le rang de la matrice des coefficients d’un modèle linéaire multivarié’, *Annales d’Économie et de Statistique* **11**, 81–111.

Gouriéroux, C., Monfort, A. and Renault, E. (1995), Inference in factor models, *in* G. S. Maddala, P. C. B. Phillips and T. N. Srinivasan, eds, ‘Advances in Econometrics and Quantitative Economics’, Blackwell, Oxford, U.K., chapter 13, pp. 311–353.

Harvey, C. R. and Siddique, A. (2000), ‘Conditional skewness in asset pricing tests’, *The Journal of Finance* **55**, 1263–1295.

Hausman, J. (1978), ‘Specification tests in econometrics’, *Econometrica* **46**, 1251–1272.

Joseph, A. S. and Kiviet, J. F. (2005), ‘Viewing the relative efficiency of IV estimators in models with lagged and instantaneous feedbacks’, *Computational Statistics and Data Analysis* **49**, 417–444.

Khalaf, L. and Kichian, M. (2005), ‘Exact tests of the stability of the Phillips curve: The Canadian case.’, *Computational Statistics and Data Analysis* **49**, 445–460.

Kiviet, J. F. and Niemczyk, J. (2007), ‘The asymptotic and finite sample distributions of OLS and simple IV in simultaneous equations’, *Computational Statistics and Data Analysis* **51**, 3296–3318.

Kraus, A. and Litzenberger, R. (1976), ‘Skewness preference and the valuation of risk assets’, *Journal of Finance* **31**, 1085–1100.

Lo, A. and MacKinlay, C. (1988), ‘Stock prices do not follow random walks: Evidence from a simple specification test’, *Review of Financial Studies* **1**, 41–66.

Shanken, J. (1986), ‘Testing portfolio efficiency when the zero-beta rate is unknown: A note’, *Journal of Finance* **41**, 269–276.

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

Simaan, Y. (1993), ‘Portfolio selection and asset pricing three parameter framework’, *Management Science* **5**, 568–577.

Smith, D. R. (2007), ‘Conditional coskewness and asset pricing’, *Journal of Empirical Finance* **41**, 91–119.

Stock, J. H., Wright, J. H. and Yogo, M. (2002), ‘A survey of weak instruments and weak identification in generalized method of moments’, *Journal of Business and Economic Statistics* **20**(4), 518–529.

Velu, R. and Zhou, G. (1999), ‘Testing multi-beta asset pricing models’, *Journal of Empirical Finance* **6**, 219–241.

Zhou, G. (1991), ‘Small sample tests of portfolio efficiency’, *Journal of Financial Economics* **30**, 165–191.

Zhou, G. (1995), 'Small sample rank tests with applications to asset pricing', *Journal of Empirical Finance* **2**, 71–93.