Finite-sample identification-robust inference for unobservable zero-beta rates and portfolio efficiency with non-Gaussian distributions *

Marie-Claude Beaulieu † Université Laval

Jean-Marie Dufour ‡ McGill University

Lynda Khalaf § Carleton University

First version: March 2002
Revised: September 2002, April 2004, September 2004, January 2005, August 2006, February 2007
This version: September 2007
Compiled: September 17, 2007, 5:56pm

*The authors thank Marine Carrasco, Christopher C. Geczy, Christian Gouriéroux, Mark Kaamstra, Raymond Kan, Shakeeb Khan, Alan Kraus, I.G. Morgan, Werner Ploberger, Carmela Quintos and Guofu Zhou for several useful comments and discussions as well as seminar participants at the Dublin 2000 EC^2 meetings, CREST (Paris), the University of British Columbia, the University of Toronto, the 2001 Canadian Economics Association meetings, CIRANO, the 2003 Winter Meetings of the Econometric Society, the 2003 conference on Simulation-Based and Finite Sample Inference in Finance (SBFSIF, Québec), the University of Rochester, and the 2004 Northern Finance Association Conference. This work was supported by the Canada Research Chair Program [Chair in Econometrics, Université de Montréal, and Chair in Environment (Environmental and Financial Econometric Analysis), Université Laval], 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 Royal Bank of Canada (Fonds Banque Royale, Université Laval), 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), and a Killam Fellowship (Canada Council for the Arts). 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.

† CIRANO and Département de finance et assurance, Université Laval. Mailing address: Département de finance et assurance, Pavillon Palasis-Prince, Université Laval, Ste-Foy, Québec, Canada G1K 7P4. 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 Holder (Environmental and Financial Econometric Analysis). Economics Department, Carleton University, CIREQ, Université de Montréal, 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

We propose exact simulation-based procedures for testing mean-variance efficiency when the zero-beta rate (gamma) is unknown and for building confidence intervals on gamma. On observing that gamma may be *weakly identified* – even though, *stricto sensu*, the problem is *not a weak instrument problem* – we propose LR-type procedures which are robust to weak identification and allow for non-Gaussian distributions. A multivariate extension of the classic Fieller method is proposed in order to build confidence sets for gamma. The exact distribution of LR-type statistics for testing efficiency is studied under both the null and alternative hypotheses. The relevant nuisance parameter structure is established and finite-sample bound procedures are proposed, which extend and improve earlier Gaussian-specific bounds. Empirical results on NYSE returns show that exact confidence sets are very different from the asymptotic ones, and allowing for non-Gaussian distributions substantially decreases the number of efficiency rejections.

Key words: capital asset pricing model; CAPM; Black; mean-variance efficiency; non-normality; weak identification; Fieller; 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.

RÉSUMÉ

Nous proposons une procédure exacte, basée sur la simulation, afin de tester l'efficience moyenne-variance d'un portefeuille lorsque le taux d'intérêt sans risque n'est pas observable. De plus, cette procédure permet de construire des intervalles de confiance pour le rendement zéro-beta qui peut êtret faiblement identifié dans le cadre de ce modèle. Les méthodes développées sont basées sur des statistiques de type ratio de vraisemblance et permettent l'utilisation d'une large classe de distributions des erreurs en plus d'être robustes à l'identification faible du zéro-beta. Une technique basée sur la méthode classique de Fieller nous sert à construire des intervalles de confiance corrigés pour la présence d'échantillons finis. Nous étudions la distribtion exacte de statistiques de type quotient de vraisemblance pour tester l'hypothèse d'efficience tant sous l'hypothèse nulle que sous la contre-hypothèse. Nous établissons la structure pertinente des paramètres de nuisance et proposons des extensions des bornes développées par Shanken pour un cadre de distributions des erreurs qui dépassent la normalité en plus que de développer une méthode plus générale menant à des bornes plus serrées que celles proposées initialement par Shanken. Les résultats empriques sur des rendements du New York Stock Exchange montrent que les intervalles de confiance exacts sont très différents des intervalles asymptotiques et que le nombre des rejets de l'hypothèse d'efficience du portefeuille de marché diminue sensiblement lorsque la possiblité d'erreurs non-gaussiennes est prise en compte.

Mots-clefs: modèle d'évaluation d'actifs financiers; CAPM; Black; efficience de portefeuille; non-normalité; identification faible; Fieller; modèle de régression multivarié; hypothèse linéaire uniforme; test exact; test de Monte Carlo; bootstrap; paramètres de nuisance; test de spécification; tests diagnostiques; GARCH; test de ratio des variances.

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

SUMMARY

We consider the problem of testing portfolio efficiency when the zero-beta rate is unknown [Black Capital Asset Pricing Model (BCAPM)]. Despite the apparently simple structure of the underlying statistical framework [a multivariate linear regression (MLR)], it is well known that standard asymptotically justified tests and confidence intervals are quite unreliable in this setup. We point out that this feature is associated with the fact that the zero-beta rate may be interpreted as a structural parameter that may be *weakly identified*, leading to a breakdown of standard asymptotic procedures based on estimated standard errors, heavy dependence on unknown nuisance parameters, and the need to use statistical methods with substantially improved finite-sample properties. This also entails that bound procedures may be required in this setup and provides a strong general motivation for developing finite-sample bounds as originally suggested by Shanken (1986, Journal of Financial Economics). The available exact procedures for the BCAPM, however, rely heavily on the assumption that model disturbances follow a Gaussian distribution, which does not appear to be satisfied by many financial return series. In this paper, we propose exact simulation-based procedures for testing mean-variance efficiency of the market portfolio and building confidence intervals for the unknown zero-beta rate. The proposed methods are based on likelihood-ratio-type statistics, allow for a wide class of error distributions (possibly heavy-tailed) and are robust to weak identification of the zero-beta rate. The exact distribution of LR-type statistics for CAPM restrictions is derived and extensions of Shanken's bounds are provided, which remain provably valid for non-Gaussian disturbance distributions. Further, we suggest a general method which yields tighter bounds in both Gaussian and non-Gaussian cases. In order to build confidence intervals for the zero-beta rate in finite samples, a technique based on generalizations (applicable in MLR models) of the classic Fieller method (for the ratio of two parameters) is proposed. The methodology described allows one to cast evidence on whether the normality assumption is too restrictive for testing the mean-variance efficiency of the benchmark portfolio when the riskless rate is unobservable. The methods suggested are applied to monthly returns on 12 portfolios of the New York Stock Exchange (NYSE) firms over the period 1925-1995 divided in five-year subperiods. The results obtained show the following: (i) multivariate normality is rejected in most subperiods, (ii) multivariate residual checks reveal no significant departures from the *i.i.d.* assumption, (iii) the exact confidence sets for the zero-beta rate are very different from the asymptotic one, and (iv) using finite-sample corrections and allowing for non-normal distributions substantially increases the proportion of cases where efficiency is accepted (more than two-thirds of the subperiods considered), as opposed to results based on asymptotic approximations or a Gaussian distributional assumption.

Contents

List of Assumptions, Propositions and Theorems

List of Tables

1. Introduction

One of the most important extensions of the Capital Asset Pricing Model (CAPM) consists in allowing for the absence of a risk-free asset. From a theoretical viewpoint, this can be due to restrictions on borrowing [Black (1972)] or an investor's riskless borrowing rate that exceeds the Treasury bill rate [Brennan (1971)]. In this case, portfolio mean-variance efficiency is defined using the expected return in excess of the zero-beta portfolio. The latter is unobservable and must be estimated in empirical applications. This leads to considerable empirical difficulties, for example in the important problem of testing mean-variance efficiency. In particular, as emphasized in the early work of Shanken (1985, 1986), large-sample approximations to the distributions of test statistics can be very unreliable in this setup, which provides a strong motivation for using econometric methods with a finite-sample distributional theory.

There are two basic approaches to estimating and assessing this version of the CAPM [denoted below as BCAPM]. The first one uses a "two-pass" approach: the alphas and betas are first estimated from time series regressions for each security, and then the zero-beta rate is estimated by a cross-sectional regression on the first-stage parameter estimates; see Black, Jensen and Scholes (1972), Fama and MacBeth (1973), and Shanken (1992). This is computationally simple but raises errors-in-variables or "generated regressors" problems [see Pagan (1984)] that can make statistical inference complex or unreliable in both finite and large samples.[1] The second approach consists in dealing jointly with the unobservable zero-beta rate, using as statistical framework a multivariate linear regression (MLR). The MLR is one of the most basic model of multivariate statistical theory and has given rise to an extensive literature, especially when errors follow a Gaussian distribution.[2] Following the seminal contributions of MacBeth (1979) and Gibbons (1982), this literature has been applied and adapted to the BCAPM by several authors; see Jobson and Korkie (1982), Kandel (1984, 1986), Amsler and Schmidt (1985), Shanken (1985, 1986), Zhou (1991), Shanken (1992), Stewart (1997), Velu and Zhou (1999), and Chou (2000).[3]

In principle, the MLR-based approach allows one to avoid the errors in variables associated with two-pass regressions. For that reason, we shall focus here on this approach and consider two basic problems: (1) testing portfolio efficiency; (2) making reliable statistical inference on the unobservable zero-beta rate, such as building a confidence interval for the zero-beta rate.

Within the BCAPM framework, testing mean-variance efficiency involves nonlinear constraints on regression coefficients. Despite these nonlinearities, the test problem is computationally simple because non-iterative analytical formulae for Gaussian likelihood-ratio (LR) test statistics are available; see Kandel (1984), Shanken (1986, 1992), Velu and Zhou (1999), and Chou (2000). It is however difficult to find reliable critical points in this context. While Gibbons (1982) used an asymptotic chi-square critical value, subsequent authors found this could lead to very serious over-rejection rates in significance tests with sample sizes commonly encountered in empirical applications [see Jobson and Korkie (1982), Shanken (1985, 1996), Campbell et al. (1997)], and various

[1]For further discussion of the related statistical problems, see Litzenberger and Ramaswamy (1979), Banz (1981), Roll (1985), Chen, Roll and Ross (1986), Shanken (1992), and Kim (1995).

[2]See, for example, Rao (1973, Chapter 8), Krishnaiah (1980), Kariya (1985), Seber (1984) and Anderson (2003).

[3]For reviews and related references, see also Kandel and Stambaugh (1989), Jagannathan and McGrattan (1995), Shanken (1996), Campbell, Lo and MacKinlay (1997), Fama and French (2004) and Perold (2004).

finite-sample corrections – such as bounds – were suggested [see Shanken (1985, 1986), Zhou (1991) and Velu and Zhou (1999)]. The finite-sample results presented by these authors depend crucially on a normality assumption, and it is not clear how they can be extended to non-Gaussian distributions. Of course, the normality assumption may be quite inappropriate in financial data used to assess the CAPM [see Richardson and Smith (1993) and Dufour, Khalaf and Beaulieu (2003)]. Besides, from the empirical viewpoint, the literature on testing efficiency within the BCAPM setup remains scarce and has produced conflicting results.[4]

In general MLR-based tests, discrepancies between asymptotic and finite sample distributions are associated with dimensionality problems: as the numbers of equations and parameters increase, the number of observations required for large-sample approximations to be reliable grows rapidly; see Shanken (1996), Campbell et al. (1997), and Dufour and Khalaf (2002). In particular, even though the asymptotic null distributions of the test statistics may be free of nuisance parameters, error cross-correlations – whose number may be quite large in systems with many portfolios – can still affect the distributions in finite samples, leading to reductions in degrees of freedom and to size distortions. Further, in the BCAPM case, the model involves nonlinear restrictions, and *dimensionality* problems compound with *identification* issues – a feature which has not apparently been discussed in the earlier literature on the BCAPM. The latter relate to the unobserved *zero-beta* rate, which is defined as the ratio of each *alpha* (intercept) to one minus *beta*. Obviously, this involves a discontinuous transformation of MLR coefficients, which becomes ill-behaved as the *betas* approach one. More precisely, the zero-beta rate is a parameter which is not identified over the whole parameter space: it is *locally almost unidentified* or (equivalently) *weakly identified* over a subset of the parameter space. Further, it is well known that such situations can strongly affect the distributions of estimators and test statistics, leading to the failure of standard asymptotic approximations, especially for Wald-type statistics [such as those used by Chou (2000)]; see, for example, Dufour (1997, 2003), Staiger and Stock (1997), Wang and Zivot (1998), Zivot, Startz and Nelson (1998), Dufour and Jasiak (2001), Kleibergen (2002), Stock, Wright and Yogo (2002) and Moreira (2003). These results should not be taken lightly in BCAPM contexts, since estimated *betas* are typically close to one (the discontinuity boundary); see Fama and MacBeth (1973).

To overcome weak identification, the theory and practice of econometrics has recently evolved in two main directions: (i) refinements in asymptotic analysis which formally account for local under-identification; examples include the local-to-zero or local-to-unity frameworks; (ii) *identification-robust* methods which control statistical error regardless of identification.[5] In particular, the second approach has the interesting feature of being able to yield exact tests and confidence sets (CS's) for structural models – which does not appear to be the case for the first approach. For this reason we shall focus here on identification-robust methods for which it is possible to derive a usable finite-sample distributional theory.

For the basic problem of testing efficiency of the market portfolio, at least three exact bound procedures have been proposed for the Gaussian case. The first one is due to Shanken (1986) and

[4]For a summary of these results, see Section 8 below.

[5]See the reviews of Stock et al. (2002) and Dufour (2003). These authors use the term "robust to weak instruments" to designate procedures whose validity is not affected by a set of instruments that does not allow one to identify structural parameters. Since we consider here a setup where instrumental variables are not explicitly required, we shall employ the term "identification-robust" which appears sufficiently general to cover the kind of situation studied in this paper.

has recently been the subject of renewed attention; see Campbell et al. (1997, chapter 5) and Chou (2000). The second bound which is developed in Stewart (1997) is directly based on a MLR extension of a general bound procedure proposed in Dufour (1989). The third one by Zhou (1991, 1995) and Velu and Zhou (1999) provides improvements in tightness by exploiting statistical results on dimensionality tests in Gaussian multivariate regressions. However, empirical evidence of multivariate non-normality of financial returns leads one to question the usefulness of such Gaussian-based exact tests.[6] In Beaulieu, Dufour and Khalaf (2006), we studied this problem for mean-variance efficiency tests in the context of the CAPM assuming observable risk-free rates. To do this, we extended Gibbons, Ross and Shanken's (1989) exact test to non-normal settings consistent with the financial theory of the CAPM and showed that allowing for elliptical non-Gaussian distributions yields a better fit in the sense that mean-variance efficiency is rejected less often than under normality.

This paper develops a general finite-sample framework for inference on the BCAPM. *First*, we propose finite-sample tests and CS's for the *zero-beta* rate. To emphasize the need for improved statistical procedures, we conduct a small simulation study on the confidence interval proposed in Campbell et al. (1997), which shows very serious level distortions. In contrast, we propose an exact CS based on "inverting" exact tests for specific values of the zero-beta rate, *i.e.* it corresponds to the set of values not rejected by this exact test. This method, which guarantees level control by construction, is a generalization of the classical confidence procedure proposed by Fieller (1954); see Dufour (1997, section 5.1). Its validity follows from the fact that the test statistic used is a proper pivotal function, *i.e.* its null distribution does not depend on nuisance parameters. In view of the recent literature on statistical inference in the presence of weak identification, this type of procedure constitutes the correct substitute to (typically unreliable) asymptotic Wald-type confidence intervals; see Dufour (1997), Staiger and Stock (1997) and Wang and Zivot (1998). In particular, in the Gaussian case, we show that the CS for the zero-beta rate can be obtained by solving a quadratic inequation – yielding a *multivariate extension* of the classical Fieller method – while for non-Gaussian distributions we propose a simulation-based extension of the latter procedure which can be implemented by numerical methods. Further, we show that the confidence intervals so obtained provide relevant information on whether efficiency is supported by the data, a property not shared by standard confidence intervals. This is due to the fact the CS may turn out to be empty, which takes place when all possible values of the zero-beta are rejected, *i.e.*, no parameter configuration is consistent with the data. In the normal error case, an empty CS for the *zero-beta* rate occurs when Shanken's bound test rejects efficiency. This property is interesting since the efficiency hypothesis actually defines the *zero-beta* rate. Furthermore, we can also use the exact confidence intervals to infer whether the estimated zero-beta rate is significantly different from the average real risk-free rate over each subperiod.

Second, we study the exact distribution of the Gaussian LR-type statistic for testing Black's CAPM under both the null and alternative hypotheses, again with possibly non-Gaussian distribu-

[6]For discussions of the class of return distributions compatible with the CAPM, the reader may consult Ross (1978), Chamberlain (1983), Ingersoll (1987, Chapter 4), Nielsen (1990), Allingham (1991), Berk (1997), and Dachraoui and Dionne (2003). Another possibility would consist in considering stable Paretian laws; see Samuelson (1967). However, since stable distributions other than the normal distribution do not have finite second moments, this requires replacing the variance of a portfolio by another measure of risk.

tions. We provide a general representation of this distribution as a function of model disturbances and nuisance parameters, which allows the latter to be easily simulated. By an invariance argument, we show that the latter depends on a reduced set of nuisance parameters, which are functions of regression and covariance coefficients. In particular, in the one factor case under the null hypothesis, the number of nuisance parameters does not exceed the number of equations (portfolios) plus one. In the standard Gaussian case, the number of nuisance parameters reduces to one.

Third, in view of the presence of nuisance parameters, we generalize Shanken's (1986) exact bound test beyond the Gaussian model. Both single and multi-factor models and a wide range of non-Gaussian distributions are allowed by our framework, including the family of elliptically symmetric distributions. The latter is relevant since it is theoretically compatible with BCAPM; see Ingersoll (1987). Empirically, we focus on multivariate Student-t and normal mixture distributions. Our results show that the Shanken and Stewart bounds obtain as a special case of our non-Gaussian bound. They also reveal that Shanken's and Stewart's bounds are in fact equivalent. We next propose a tighter bound, which involves a numerical simulation-based search for the tightest cut-off point. This is implemented using a maximized Monte Carlo (MMC) test procedure [Dufour (2006)]. This procedure is based on the following fundamental property: when the distribution of a test statistic depends on nuisance parameters, the desired level α is achieved by comparing the largest p-value (over all nuisance parameters consistent with the null hypothesis) with the level α of the test: a (simulated) p-value function conditional on relevant nuisance parameters is numerically maximized (with respect to these parameters), and the test is significant if the largest p-value is not larger than α.

Fourth, we formally deal with unknown distributional parameters (*e.g.* the degrees-of-freedom for the Student-t distribution, and the mixing probability and scale-ratio parameters for normal mixtures). To do this, we apply the MMC technique with the following modification [based on extending the two-step procedures proposed in Dufour (1990) and Dufour and Kiviet (1996)]: the maximization is restricted to an exact CS estimate for the relevant nuisance parameters. The latter set is obtained by inverting a distributional goodness-of-fit (GF) test. This approach allows to formally take into consideration the fact that the null hypothesis *jointly* restricts the error distribution in addition to regression coefficients. We also run exact multivariate misspecification tests (multivariate GARCH and variance ratio tests) which provide level-correct checks of the residuals for departures from the maintained *i.i.d.* hypothesis.

Fifth, the methodology developed is used to examine efficiency of the market portfolio for monthly returns on New York Stock Exchange (NYSE) portfolios, built from the University of Chicago Center for Research in Security Prices (CRSP) 1926-1995 data base. Results based on asymptotic distributional theory, exact tests under normality as well as exact p-values for non-normal elliptical distributions are presented. We find that mean-variance efficiency of the market portfolio in the context of BCAPM is rejected over a smaller number of subperiods under the Student and the normal mixture distributions than under the normal. Finally, using exact CS's for γ, we observe that they do not lead to the same decision as the asymptotic confidence intervals when checking whether the average risk-free rate is covered by the CS for γ.

The paper is organized as follows. Section 2 sets the framework and the basic test statistics that will be considered. In Section 3, we discuss the identification issues associated with the unobservable zero-beta rate and present a small simulation study illustrating the statistical difficulties entailed by this feature. In Section 4, we propose finite-sample tests for specific values of the zero-beta rate, and the corresponding exact CS are derived in Section 5. The exact distribution of a LR-type statistic for testing mean-variance efficiency is established in Section 6, and bound procedures are proposed in Section 7. In Section 8, we report the empirical results. We conclude in Section 9.

2. Framework and motivation for exact inference

Let R_{it}, $i = 1, \ldots, n$, be the returns on n securities in period t, and $\tilde{R}_{Mt}$ the return on a market benchmark $(t = 1, \ldots, T)$. Our analysis of the BCAPM model is based on the following standard MLR setup [Gibbons (1982), Shanken (1986), MacKinlay (1987)]:

$$R_{it} = a_i + \beta_i \tilde{R}_{Mt} + u_{it}, \quad i = 1, \ldots, n, \ t = 1, \ldots, T, \tag{2.1}$$

where the u_{it} are random disturbances. We assume we can condition on $\tilde{R}_M = \big(\tilde{R}_{M1}, \ldots, \tilde{R}_{MT}\big)'$, *i.e.* we can take $\tilde{R}_M$ as fixed for statistical inference. Furthermore, the vectors $V_t = (u_{1t}, \ldots, u_{nt})'$ have the form:

$$V_t = K'W_t, \quad t = 1, \ldots, T, \tag{2.2}$$

where K' is an unknown non-singular matrix and the distribution of $W = [W_1, \ldots, W_T]'$ is: either (i) fully specified, or (ii) specified up to an unknown *distributional shape parameter* ν. Although we present general results which require *no further regularity assumptions* on the error terms, the CAPM imposes distributional restrictions, which entail that the distribution of W belongs to a specific distributional family $\mathcal{H}_W(\mathcal{D}, \nu)$, where $\mathcal{D}$ represents a distribution type and $\nu \in \Omega_{\mathcal{D}}$ any (eventual) nuisance parameter characterizing the distribution. Below, we focus on multivariate normal $(\mathcal{D}_N)$, Student-t $(\mathcal{D}_t)$ and normal mixture $(\mathcal{D}_m)$ distributions for W:

$$\mathcal{H}_W(\mathcal{D}_N) : W_t \overset{i.i.d}{\sim} \mathrm{N}\big[0, I_n\big], \tag{2.3}$$

$$\mathcal{H}_W(\mathcal{D}_t, \kappa) : W_t = Z_{1t}/(Z_{2t}/\kappa)^{1/2}, \ Z_{1t} \overset{i.i.d}{\sim} \mathrm{N}\big[0, I_n\big], \ Z_{2t} \overset{i.i.d}{\sim} \chi^2(\kappa), \tag{2.4}$$

$$\mathcal{H}_W(\mathcal{D}_m, \pi, \omega) : W_t = I_t(\pi)Z_{1t} + [1 - I_t(\pi)]Z_{3t}, \ Z_{3t} \overset{i.i.d}{\sim} \mathrm{N}\big[0, \omega I_n\big], \ 0 < \pi < 1, \tag{2.5}$$

where Z_{2t} and Z_{3t} are independent of Z_{1t}, and $I_t(\pi)$ is an indicator variable independent of $\big(Z_{1t}, Z_{3t}\big)$ such that $\mathsf{P}[I_t(\pi) = 0] = 1 - \mathsf{P}[I_t(\pi) = 1] = \pi$. So, in (2.2), $\nu = \kappa$ under (2.4), and $\nu = (\pi, \omega)$ under (2.5). Clearly, if $E(W_t W_t') = I_n$, the covariance matrix of W_t is $\Sigma = K'K$. These distribution families are consistent with the theoretical foundations of the CAPM [see Ingersoll (1987)] and are also empirically relevant since they allow for the "fat tails" typically observed in financial returns [see Fama (1965)]. Note that Σ is positive definite and is not otherwise restricted, but further constraints on K may be needed in order to determine K uniquely. For example, if W_t is Gaussian, we may assume that K is upper triangular so that $\Sigma = K'K$ corresponds to the Cholesky factorization of Σ. So we assume that:

$$K \in \mathcal{J} \tag{2.6}$$

where $\mathcal{J}$ is a multiplicative group of $n \times n$ matrices. For example, $\mathcal{J}$ may be the set of all $n \times n$ non-singular matrices or the subgroup of non-singular lower-triangular matrices.

The testable implication of the BCAPM on (2.1) is the following one: there is a scalar γ, the return on the zero-beta portfolio, such that

$$\mathcal{H}_{\mathrm{B}} : a_i - \gamma_0(1 - \beta_i) = 0\,, \quad i = 1,\, \ldots\,,\, n, \quad \text{for some } \gamma_0 \in \Gamma\,, \tag{2.7}$$

where Γ is some set of "admissible" values for γ, such as $\Gamma = \mathbb{R}$ (unrestricted), $\Gamma = [0, \infty)$ or $\Gamma = [-1, 1]$. Since γ is unknown, $\mathcal{H}_{\mathrm{B}}$ is nonlinear. However, the more restrictive hypothesis

$$\mathcal{H}(\gamma_0) : a_i - \gamma_0(1 - \beta_i) = 0\,,\ i = 1,\, \ldots\,,\, n, \quad \text{(where } \gamma_0 \text{ is specified)} \tag{2.8}$$

is linear. This observation underlies our exact bound CAPM test and CS for γ.

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

$$Y = XB + U\,, \quad U = WK\,, \tag{2.9}$$

where $Y = [Y_1,\, \ldots\,,\, Y_n]$ is a $T \times n$ matrix, X is $T \times k$ matrix of rank k, $U = [U_1,\, \ldots\,,\, U_n] = [V_1,\, \ldots\,,\, V_T]'$ is a $T \times n$ matrix of error terms which are independent of X. For model (2.1), we have: $Y = [R_1,\, \ldots\,,\, R_n]\,,\, X = \left[\iota_T,\, \tilde{R}_{\mathrm{M}}\right]\,,\ R_i = (R_{i1},\, \ldots\,,\, R_{iT})'\,,\, \tilde{R}_{\mathrm{M}} = \left(\tilde{R}_{\mathrm{M}1},\, \ldots\,,\, \tilde{R}_{\mathrm{M}T}\right)'$, $B = [a, \beta]'$, $a = (a_1,\, \ldots\,,\, a_n)'$ and $\beta = (\beta_1,\, \ldots\,,\, \beta_n)'$. In this context, one of the most common way of testing $\mathcal{H}_{\mathrm{B}}$ consists in maximizing the Gaussian log-likelihood

$$\ln[L(Y,\, B\,,\, \Sigma)] = -\frac{nT}{2}(2\pi) - \frac{T}{2}\ln(|\Sigma|) - \frac{1}{2}\mathrm{tr}[\Sigma^{-1}(Y - XB)'(Y - XB)] \tag{2.10}$$

under the null and the alternative hypotheses, and computing the corresponding *likelihood ratio* (LR) test criterion

$$LR_{\mathrm{B}} = T\ln(\Lambda_{\mathrm{B}}) \quad \text{with} \quad \Lambda_{\mathrm{B}} = |\hat{\Sigma}_{\mathrm{B}}|/|\hat{\Sigma}|\,, \tag{2.11}$$

where $\hat{\Sigma}_{\mathrm{B}}$ is the maximum likelihood (ML) estimator of Σ under $\mathcal{H}_{\mathrm{B}}$, $\hat{\Sigma} = \hat{U}'\hat{U}/T$, $\hat{U} = Y - X\hat{B}$, $\hat{B} = (X'X)^{-1}X'Y$. Similarly, the LR statistic to test $\mathcal{H}(\gamma_0)$ is

$$LR(\gamma_0) = T\ln[\Lambda(\gamma_0)] \quad \text{with} \quad \Lambda(\gamma_0) = |\hat{\Sigma}(\gamma_0)|/|\hat{\Sigma}| \tag{2.12}$$

where $\hat{\Sigma}(\gamma_0)$ is the restricted ML estimator of Σ under $\mathcal{H}(\gamma_0)$. It is easy to see that the two statistics are related as follows:

$$|\hat{\Sigma}_{\mathrm{B}}| = \inf\{|\hat{\Sigma}(\gamma_0)| : \gamma_0 \in \Gamma\}\,, \quad LR_{\mathrm{B}} = \inf\{LR(\gamma_0) : \gamma_0 \in \Gamma\} = LR(\hat{\gamma})\,, \tag{2.13}$$

where $\hat{\gamma}$ is the QML estimator of γ; see Shanken (1986).

Because the errors may not follow Gaussian distributions in our setup, we shall call LR_{B} and $LR(\gamma_0)$ *quasi likelihood ratio* (QLR) criteria and the ML estimators on which they are based *Gaussian quasi maximum likelihood* (QML) estimators. Note also that the condition $\gamma_0 \in \Gamma$ may involve inequality or nonlinear constraints that can lead to non-regular distributions. The simulation-based approach we propose below deals in a transparent way with such difficulties. For further reference, we will denote the observed value of these statistics (calculated from the available sample) as $LR^{(0)}_{\mathrm{B}}$ and $LR^{(0)}(\gamma_0)$, respectively. $\mathsf{P}_{(B,\,K)}$ represents the distribution of Y when the parameters are $(B,\,K)$, while for any matrix A, the symbol $M(A)$ denotes the projection matrix $M(A) = I - A(A'A)^{-}A'$.

3. The zero-beta rate as a weakly identified parameter

Although several computationally simple procedures have been developed to obtain constrained QML estimates [see Gibbons (1982), Shanken (1986), Zhou (1991), and Chou (2000)], CS estimates of the key parameter γ appear to be lacking. A Wald-type QML formula for an asymptotic standard error is provided by Campbell et al. (1997, Chapter 5, equation 5.3.81) for the single *beta* case:

$$\mathsf{Var}(\hat{\gamma}) = \frac{1}{T}\left(1 + \frac{(\bar{R}_{\mathrm{M}} - \hat{\gamma})^2}{\varDelta^2_{\mathrm{M}}}\right)[(\iota_n - \hat{\beta})'\hat{\varSigma}^{-1}(\iota_n - \hat{\beta})]^{-1} \tag{3.1}$$

where $\hat{\beta}$ is an $n \times 1$ vector which includes the unconstrained OLS estimates of $\beta_i\,,\ i = 1,\,\ldots\,,\,n,\ \iota_n = (1,\,\ldots\,,\,1)'$, $\bar{R}_{\mathrm{M}}$ and $\varDelta^2_{\mathrm{M}}$ are respectively the sample mean and variance of the benchmark returns (on the right-hand-side), and $\hat{\gamma}$ is the QML estimator of γ (see Section 7.1 for further discussion). It seems however that this expression has not often been used in practice.

A central difficulty with model (2.1) jointly with (2.7) stems from the fact that, even though a_i and β_i are well identified (because they can be interpreted as regression coefficients), the zero-beta rate γ is defined through a nonlinear transformation that may fail to be well-defined. Namely, the ratio $\gamma = a_i/(1-\beta_i)$ is not defined when $\beta_i = 1$ or, equivalently, the equation $a_i = \gamma(1-\beta_i)$ does not have a unique solution when $\beta_i = 1$. In other words, γ is *not identifiable* when $\beta_i = 1\,,\ i = 1,\,\ldots\,,\,n.$ So γ may be quite difficult to estimate with reasonable precision if the coefficients β_i are close to one. γ is a *locally almost unidentified* (LAU) parameter [Dufour (1997)] or, following the weak-instrument literature, γ is a *weakly identified* parameter on some subset of the parameter space.

In such situations, many standard test statistics, such as Wald-type statistics, cannot be pivotal irrespective of the sample size (*i.e.*, their distributions depend on unknown nuisance parameters), so that the corresponding tests can be arbitrarily unreliable and CS's based on such statistics become invalid in a fundamental way (*e.g.*, the true size of a CS with nominal level 0.95 is in fact equal to zero). In particular, asymptotic standard errors become extremely unreliable and standard asymptotically justified t-type tests and confidence intervals have sizes that may deviate arbitrarily from their nominal levels. Both the finite and large-sample distributional theory of most test statistics can be heavily affected, so that important adjustments are typically required. These issues have been explored in the recent literature on weak instruments; see, for example, Staiger and Stock (1997), Wang and Zivot (1998), Zivot et al. (1998), Dufour and Jasiak (2001), Stock and Wright (2000), Kleibergen (2002), Moreira (2003), and the reviews of Stock et al. (2002) and Dufour (2003).

Even though this work focuses on inference in simultaneous equations models (or IV regressions), it is important to note that the problem we consider here is fundamentally the same. To see that the issue can matter in the context of inference on γ that may be weakly identified, it will be

Table 1. Asymptotic confidence interval for the zero-beta rate: coverage probabilities

Parameters of	(2)	(3)	(4)	(5)
Sample	$T = 60$	$T = 100$	$T = 500$	$T = 1000$
1927-30	0.083	0.117	0.304	0.411
1931-35	0.036	0.054	0.114	0.209
1926-40	0.062	0.078	0.158	0.296
1941-45	0.071	0.124	0.335	0.432
1946-50	0.111	0.177	0.445	0.589
1951-55	0.157	0.236	0.543	0.669
1956-60	0.138	0.267	0.546	0.652
1961-65	0.075	0.127	0.272	0.426
1966-70	0.085	0.139	0.392	0.546
1971-75	0.064	0.090	0.242	0.312
1976-80	0.072	0.121	0.390	0.508
1981-85	0.099	0.139	0.376	0.525
1986-90	0.101	0.125	0.301	0.421
1991-95	0.162	0.187	0.302	0.360

Note – Numbers shown are estimated coverage probabilities for an asymptotic confidence interval for γ based on (3.1) with nominal level 0.95, using 1000 simulated samples conformable with (2.1). These are drawn using estimated parameters (including error covariances), from the empirical model analyzed in Section 8, imposing efficiency. The regressor matrix includes, in addition to a constant, a simulated variate drawn as uniformly distributed over the observed range of benchmark returns. Empirical coverage probabilities are the proportion of cases where the true value of γ falls within the estimated confidence interval.

useful to look at the results of a small simulation experiment. Using estimated parameters from the empirical model estimated in Section 8, including the error covariances, we generated 1000 realizations of the MLR (2.1) under test, imposing efficiency. For each simulated sample, a Wald-type confidence interval with nominal level of 95% was computed using (3.1) as the standard error, and the empirical proportion of cases where the true value of γ fell in the confidence interval was evaluated. The regressor matrix includes, in addition to a constant, a simulated variate drawn as uniformly distributed over the observed range of benchmark returns, to allow the analysis of several sample sizes. The results are reported in Table 1. The coverage probabilities obtained can be severely low: instead of being around 95%, they range from 3.6% to just 16% with observed sample sizes, from 5.6% to just 18% with $T = 100$, and from 20% to 67% when $T = 1000$. These results motivate the finite-sample approach we propose in what follows.

4. Exact identification-robust tests for the zero-beta rate

We will first derive the exact null distribution of the QLR statistic $LR(\gamma_0)$ under $\mathcal{H}(\gamma_0)$, where γ_0 is known. This will allow us to build a CS for γ and yield a way of testing efficiency. The basic distributional result for that purpose is given by the following theorem.

Theorem 4.1 DISTRIBUTION OF THE MEAN-VARIANCE CAPM TEST FOR A KNOWN ZERO-BETA RATE. *Under* (2.1), (2.2) *and* (2.8), *the statistic* $LR(\gamma_0)$ *in* (2.12) *is distributed like*

$$\overline{LR}(\gamma_0, W) = T\ \ln(\left|W'\bar{M}(\gamma_0)W\right| / \left|W'MW\right|) \tag{4.1}$$

where $\bar{M}(\gamma_0) = M + X(X'X)^{-1}H(\gamma_0)'[H(\gamma_0)(X'X)^{-1}H(\gamma_0)']^{-1}H(\gamma_0)(X'X)^{-1}X'$ *and* $H(\gamma_0)$ *is the row vector* $(1, \gamma_0)$.

Proofs are given in the Appendix. In the Gaussian case (2.3), we have:

$$[(T-1-n)/n][\Lambda(\gamma_0)-1] \sim F(n, T-1-n)\,; \tag{4.2}$$

see Dufour and Khalaf (2002, Appendix, Theorem A.1).[7] This result was exploited by Gibbons et al. (1989) in studying mean-variance efficiency with observable risk-free rate. Indeed, testing $\mathcal{H}(\gamma_0)$ is equivalent to testing whether the regression intercepts are jointly zero in a market model where returns are in excess of γ_0: $\mathcal{H}(\gamma_0) \Leftrightarrow \bar{a}_i = 0\,,\ i = 1, \ldots, n$, in the context of $R_{it} - \gamma_0 = \bar{a}_i + \beta_i(\tilde{R}_{\mathrm{M}t} - \gamma_0) + u_{it}\,, i = 1, \ldots, n,\ t = 1, \ldots, T.$

For non-Gaussian distributions compatible with (2.2), Theorem **4.1** shows that the exact distribution of $LR(\gamma_0)$, although non-standard, may easily be simulated once the matrix X, the distribution of W and the value γ_0 [given by $\mathcal{H}(\gamma_0)$] are set. So the MC test method can be easily applied [see Dufour (2006)]. In general, the MC test method assesses the rank of the observed value of a test statistic $\big[$denoted $S^{(0)}\big]$, relative to a finite number N of simulated statistics [denoted $S^{(1)}, \ldots, S^{(N)}$] drawn under the null hypothesis. Conforming with (2.2), we assume that $S^{(1)}, \ldots, S^{(N)}$ can be simulated given: (i) a value of ν, (ii) N draws $W^{(1)}, \ldots, W^{(N)}$, from the distribution of W [which under (2.2) can be simulated once ν is specified], (iii) a vector of parameters (denoted η) which affect the distribution of the test statistic, and (iv) the test function $\bar{S}(\eta, W)$ which depends on η, W and X.[8] In other words, on drawing N samples from the distribution of W (which depends on ν) and computing $\bar{S}(\eta, W)$ for each simulated sample, we get the vector $\bar{S}_N(\eta, \nu) = \big[\bar{S}\big(\eta, W^{(1)}\big), \ldots, \bar{S}\big(\eta, W^{(N)}\big)\big]'$. Then a *MC* p-value is defined as:

$$p_N[S^{(0)}|\bar{S}_N(\eta, \nu)] = \frac{NG_N\big[S^{(0)}\,;\, \bar{S}_N(\eta, \nu)\big] + 1}{N+1}\,, \tag{4.3}$$

$$G_N\big[S^{(0)}\,;\, \bar{S}_N(\eta, \nu)\big] = \frac{1}{N}\sum_{j=1}^{N} I_{[0,\infty)}\big[\bar{S}(W^{(j)}, \eta) - S^{(0)}\big]\,, \tag{4.4}$$

where $I_A[x] = 1$, if $x \in A$, and $I_A[x] = 0$, if $x \notin A$. In the case of $LR(\gamma_0)$, we have $S^{(0)} \equiv LR^{(0)}(\gamma_0)$, $\eta \equiv \gamma_0$, $\bar{S}\big(\eta, W^{(i)}\big) = \overline{LR}\big(\gamma_0, W^{(i)}\big)$, $i = 1, \ldots, N$, hence

$$\hat{p}_N(\gamma_0, \nu) \equiv p_N\big[LR^{(0)}(\gamma_0)\big|\overline{LR}_N(\gamma_0, \nu)\big] \tag{4.5}$$

[7] A summary of these results appears in Appendix A.

[8] For notational simplicity – in view of the fact that X is taken as fixed – the dependence upon X is implicit through the definition of the $\bar{S}$ mapping.

where $\overline{LR}_N(\gamma_0,\, \nu) = \big[\overline{LR}\big(\gamma_0,\, W^{(1)}\big),\, \ldots\,,\, \overline{LR}\big(\gamma_0,\, W^{(N)}\big)\big]'$. As a result of Theorem **4.1**, we have, under the null hypothesis $\mathcal{H}(\gamma_0)$ in conjunction with $\mathcal{H}_W(\mathcal{D},\, \nu)$:

$$\mathsf{P}\big[\hat{p}_N(\gamma_0,\, \nu_0) \leq \alpha\big] = \alpha\,, \text{when } \nu = \nu_0, \tag{4.6}$$

$$\mathsf{P}\big[\sup\{\hat{p}_N(\gamma_0,\, \nu_0) : \nu_0 \in \Omega_{\mathcal{D}}\} \leq \alpha\big] \leq \alpha\,, \text{when } \nu \text{ may be unknown.} \tag{4.7}$$

For further reference, we will call $\hat{p}_N(\gamma_0,\, \nu)$ a *pivotal MC* (PMC) p-value, to emphasize that it is based on the null distribution of a pivotal (in the sense of invariance to B and K) statistic. These results provide the basis for the CS procedure we propose in what follows for inference on the zero-beta rate.[9]

5. Exact confidence sets for the zero-beta rate

The problem of building a CS for γ (given the definition of this parameter) can be viewed as an extension of the classical "Fieller problem" which focuses on estimating the ratio of the means of two normal samples with the same variance (and no regressors). Fieller (1940, 1954) gave a finite-sample solution to this problem, which can be interpreted as the inversion of a LR test and avoids the non-regularities associated with the fact that the denominator could be equal or close to zero. While it is relatively straightforward to extend Fieller's solution to univariate linear regressions [see Zerbe (1978) and Dufour (1997)], the MLR model here is much more complex. In this section, we provide the required extension, which is based on "inverting" finite-sample simulation-based LR-type tests. Specifically, we consider "inverting" the LR test defined in Theorem **4.1**, rather than standard t or F tests in a single regression - as one in the earlier literature on this topic.[10] Our analysis reveals crucial differences between the Gaussian and non-Gaussian cases.

5.1. Fieller-type confidence sets for zero-beta rate: Gaussian case

Consider the Gaussian one-factor model (2.1) - (2.3). Let us denote by $\hat{a} = (\hat{a}_1,\, \ldots\,,\, \hat{a}_n)'$ and $\hat{\beta} = (\hat{\beta}_1,\, \ldots\,,\, \hat{\beta}_n)'$ the unrestricted least squares estimators, and set $\hat{\delta} = \iota_n - \hat{\beta}$, $\hat{\mu}_M = \frac{1}{T}\sum_{t=1}^{T} \tilde{R}_{\mathrm{M}t}$, $\hat{\sigma}_M^2 = \frac{1}{T}\sum_{t=1}^{T}(\tilde{R}_{\mathrm{M}t} - \hat{\mu}_M)^2$. Under $\mathcal{H}_{\mathrm{B}}$, the ratios $a_i/(1-\beta_i)\,,\, 1,\, \ldots\,,\, n$, are equal. We wish to build a CS for the coefficient γ such that $\gamma = a_i/(1-\beta_i)\,,\, 1,\, \ldots\,,\, n$. To do so, we consider first the problem of testing whether γ has a given value γ_0, *i.e.* $\mathcal{H}(\gamma_0)$ as defined in (2.8). In this case, the LR statistic is a monotonic transformation of the Wald-type (or Hotelling) statistic

$$\mathcal{W}(\gamma_0) = \frac{T-n-1}{n}\,\frac{\big(\hat{a} - \hat{\delta}\gamma_0\big)'\hat{\Sigma}^{-1}\big(\hat{a} - \hat{\delta}\gamma_0\big)}{1 + \big[(\hat{\mu}_M - \gamma_0)^2/\hat{\sigma}_M^2\big]}, \tag{5.1}$$

whose null distribution does not depend on γ_0, so that the same critical point can be used to test any value γ_0. Indeed, under $\mathcal{H}_0(\gamma_0)$, $\mathcal{W}(\gamma_0)$ follows a Fisher distribution $F\,(n, T-n-1)$; see

[9] Further discussion of exact MC p-values is provided in Appendix C.

[10] See Fieller (1940, 1954), Koschat (1987), Zerbe (1978), Buonaccorsi and Iyer (1984) and Dufour (1997, Section 5.1).

Campbell et al. (1997, equation 5.3.80, page 202). Let F_α denote the cut-off point for a test with level α based on the $F(n, T-n-1)$ distribution, *i.e.*, $\mathsf{P}[F(n, T-n-1) \geq F_\alpha] = \alpha$. Then the set

$$CF_\gamma(\alpha) = \{\gamma_0 \in \Gamma : \mathcal{W}(\gamma_0) \leq F_\alpha\} \tag{5.2}$$

has level $1-\alpha$ for γ. In other words, the probability that γ be covered by $CF_\gamma(\alpha)$ is not smaller than $1-\alpha : \mathsf{P}[\gamma \in CF_\gamma(\alpha)] \geq 1-\alpha$.

Indeed, the equality $\mathsf{P}[\gamma \in CF_\gamma(\alpha)] = 1-\alpha$ holds here. On noting that the inequality $\mathcal{W}(\gamma_0) \leq F_\alpha$ can be rewritten as

$$M_F(\gamma_0) - \frac{nF_\alpha}{T-n-1} N_F(\gamma_0) \leq 0\,, \tag{5.3}$$

$$M_F(\gamma_0) = \big(\hat{a} - \hat{\delta}\gamma_0\big)'\hat{\Sigma}^{-1}\big(\hat{a} - \hat{\delta}\gamma_0\big) = \big(\hat{\delta}'\hat{\Sigma}^{-1}\hat{\delta}\big)\gamma_0^2 - \big(2\hat{\delta}'\hat{\Sigma}^{-1}\hat{a}\big)\gamma_0 + \hat{a}'\hat{\Sigma}^{-1}\hat{a}\,, \tag{5.4}$$

$$N_F(\gamma_0) = 1 + \frac{(\hat{\mu}_M - \gamma_0)^2}{\hat{\sigma}_M^2} = \frac{1}{\hat{\sigma}_M^2}\gamma_0^2 - \frac{2\hat{\mu}_m}{\hat{\sigma}_M^2}\gamma_0 + 1 + \frac{\hat{\mu}_M^2}{\hat{\sigma}_M^2}\,, \tag{5.5}$$

we see, after a few manipulations, that $CF_\gamma(\alpha)$ reduces to a simple quadratic inequation:

$$CF_\gamma(\alpha) = \{\gamma_0 \in \Gamma : A\gamma_0^2 + B\gamma_0 + C \leq 0\}\,, \tag{5.6}$$

$$A = \hat{\delta}'\hat{\Sigma}^{-1}\hat{\delta} - \left(\frac{nF_\alpha}{T-n-1}\right)\frac{1}{\hat{\sigma}_M^2}\,, \quad B = 2\left[\left(\frac{nF_\alpha}{T-n-1}\right)\frac{\hat{\mu}_M}{\hat{\sigma}_M^2} - \hat{\delta}'\hat{\Sigma}^{-1}\hat{a}\right]\,, \tag{5.7}$$

$$C = \hat{a}'\hat{\Sigma}^{-1}\hat{a} - \left(\frac{nF_\alpha}{T-n-1}\right)\left[1 + \frac{\hat{\mu}_M^2}{\hat{\sigma}_M^2}\right]\,. \tag{5.8}$$

For $\Gamma = \mathbb{R}$, the resulting CS can take several forms depending on the roots of the polynomial $A\gamma_0^2 + B\gamma_0 + C$: (a) a closed interval; (b) the union of two unbounded intervals; (c) the entire real line; (d) an empty set.[11] Case (a) corresponds to a situation where γ is well identified, while (b) and (c) correspond to unbounded CS's and indicate (partial or complete) non-identification. The possibility of getting an empty CS may appear surprising. But, on hindsight, this is quite natural: it means that no value of γ_0 does allow $\mathcal{H}(\gamma_0)$ to be acceptable. Since the efficiency hypothesis $\mathcal{H}_{\mathrm{B}}$ states that *there exists* a real scalar γ such that $a_i = (1-\beta_i)\gamma\,,\ i = 1, \ldots, n$, this can be interpreted as a rejection of $\mathcal{H}_{\mathrm{B}}$. Further, under $\mathcal{H}_{\mathrm{B}}$, the probability that $CF_\gamma(\alpha)$ covers the true value γ is $1-\alpha$, and an empty set obviously does not cover γ. Consequently, the probability that $CF_\gamma(\alpha)$ be empty $[CF_\gamma(\alpha) = \emptyset]$ cannot be greater than α under $\mathcal{H}_{\mathrm{B}} : \mathsf{P}[CF_\gamma(\alpha) = \emptyset] \leq \alpha$. The event $CF_\gamma(\alpha) = \emptyset$, is a test with level α for $\mathcal{H}_{\mathrm{B}}$ under normality.

5.2. Fieller-type confidence sets for zero-beta rate: non-Gaussian case

The quadratic CS described above relies heavily on the fact that the same critical point F_α can be used to test all values of γ_0. This occurs under the Gaussian distributional assumption, but not necessarily if errors are not Gaussian. Although the quadratic CS will remain "asymptotically valid"

[11] Further discussion of such quadratic confidence sets is available in Dufour (1997, Section 5.1), Dufour and Jasiak (2001), Zivot et al. (1998), and Dufour and Taamouti (2005).

as long as $\mathcal{W}(\gamma)$ converges to a $\chi^2(n)$ distribution (which does not depend on nuisance parameters nor on the true parameter value of γ), this cannot provide an exact CS. So we now study how the Fieller-type procedure described above can be extended to allow for possibly non-Gaussian disturbances.

Again, the problem we face can be tackled by inverting an α-level test based on $LR(\gamma_0)$. However, the test is now performed by simulation (as a Monte Carlo test) and there is no reason to expect that the null distribution of the test statistic is the same for all values of γ_0. Consider the MC p-value $\hat{p}_N(\gamma_0,\, \nu)$ function associated with this statistic, as defined in (4.5). Since the critical region $\hat{p}_N(\gamma_0,\, \nu) \leq \alpha$ has level α for testing $\gamma = \gamma_0$ when ν is known, the set of γ_0 values for which $\hat{p}_N(\gamma_0,\, \nu)$ exceeds α, *i.e.*

$$C_\gamma(\alpha;\, \nu) = \left\{\gamma_0 \in \varGamma : \hat{p}_N(\gamma_0,\, \nu) > \alpha\right\}, \tag{5.9}$$

is a CS with level $1-\alpha$ for γ. Similarly, when ν is not specified, the test $\sup\{p_N(\gamma_0,\, \nu_0) : \nu_0 \in \varOmega_{\mathcal{D}}\} \leq \alpha$ yields the CS:

$$C_\gamma(\alpha;\, \mathcal{D}) = \left\{\gamma_0 \in \varGamma : \sup\{\hat{p}_N(\gamma_0,\, \nu_0) : \nu_0 \in \varOmega_{\mathcal{D}}\} > \alpha\right\}, \tag{5.10}$$

whose level is also $1-\alpha$.[12] This holds for any sample size and regressor design. In contrast with the Gaussian case, there is no closed form in this case, and the set of acceptable values $C_\gamma(\alpha;\, \nu)$ or $C_\gamma(\alpha;\, \mathcal{D})$ must be drawn by numerical methods, such as a grid search. Note also that the identity $LR(\hat{\gamma}) = \inf\{LR(\gamma_0) : \gamma_0 \in \varGamma\}$ entails that $\hat{\gamma}$ must belong to the CS, provided its level is larger than zero.

It is not clear, however, that these confidence sets can only take four general shapes as occurs for $CF_\gamma(\alpha)$ in (5.2). We have no closed-form description of the structure of $C_\gamma(\alpha;\, \nu)$ or $C_\gamma(\alpha;\, \mathcal{D})$. While these can be bounded intervals (this is showed numerically in Section 8), $C_\gamma(\alpha;\, \nu)$ or $C_\gamma(\alpha;\, \mathcal{D})$ must be unbounded with a high probability if γ is not identifiable or weakly identified [see Dufour (1997)]. An empty CS is also possible and provides evidence that the efficiency hypothesis $\mathcal{H}_{\mathrm{B}}$ is not compatible with the data. The event $C_\gamma(\alpha;\, \nu) = \emptyset$ [or $C_\gamma(\alpha;\, \mathcal{D}) = \emptyset$] is a test with level α for $\mathcal{H}_{\mathrm{B}}$ under assumption (2.2).

6. Invariance and exact distribution of Black QLR

In this section, we study the exact distribution of the test statistic $LR_{\mathrm{B}} = T\ln(\varLambda_{\mathrm{B}})$ for Black's CAPM under the null hypothesis $\mathcal{H}_{\mathrm{B}}$ and the alternative MLR model. We provide a general representation of this distribution as a function W and nuisance parameters. In particular, by an invariance argument, we show that the latter depends on B and K only through a lower-dimensional function of B and K, even when the null hypothesis $\mathcal{H}_{\mathrm{B}}$ does not hold. Further, under $\mathcal{H}_{\mathrm{B}}$, the number of such free parameters is at most $n+1$.

[12] In view of this asymptotic validity of the quadratic confidence set $CF_\gamma(\alpha)$ described in Section 5.1, the latter may be used as a reasonable approximation to the exact confidence set.

To do this, we consider first the following equivalent representation of model (2.1):

$$R_{it} - \tilde{R}_{\mathrm{M}t} = a_i + \delta_i \tilde{R}_{\mathrm{M}t} + u_{it}\,, \quad i = 1, \ldots, n, \; t = 1, \ldots, T, \tag{6.1}$$

where $\delta_i = \beta_i - 1$, or in matrix form

$$\tilde{Y} = XC + U \tag{6.2}$$

where $\tilde{Y} = Y - \tilde{R}_{\mathrm{M}}\iota_n'$, $\tilde{R}_{\mathrm{M}} = \left(\tilde{R}_{\mathrm{M}1}, \ldots, \tilde{R}_{\mathrm{M}T}\right)'$, $C = B - \Delta = [a, \beta - \iota_n]'$, $\Delta = [0, \iota_n]'$, and $X = \left[\iota_T, \tilde{R}_{\mathrm{M}}\right]$. Then the null hypotheses $\mathcal{H}(\gamma_0)$ and $\mathcal{H}_{\mathrm{B}}$ can be rewritten as:

$$\tilde{\mathcal{H}}(\gamma_0) : H(\gamma_0)C = 0\,, \quad \text{(where } \gamma_0 \text{ is specified)} \tag{6.3}$$

$$\tilde{\mathcal{H}}_{\mathrm{B}} : H(\gamma_0)C = 0\,, \quad \text{for some } \gamma_0 \in \Gamma, \tag{6.4}$$

where $H(\gamma_0) = (1, \gamma_0)$. Since the Jacobian of the transformation from $\mathrm{vec}(\tilde{Y})$ to $\mathrm{vec}(Y)$ is a unit matrix, the Gaussian log-likelihood function for model (6.2) is

$$\ln[\tilde{L}(\tilde{Y}, C\,, \Sigma)] = \ln[L(Y - \tilde{R}_{\mathrm{M}}\iota_n', B - \Delta\,, \Sigma)] = \ln[L(Y, B\,, \Sigma)] \tag{6.5}$$

and the corresponding Gaussian QLR statistics for testing $\tilde{\mathcal{H}}(\gamma_0)$ and $\tilde{\mathcal{H}}_{\mathrm{B}}$ numerically coincide with $LR(\gamma_0)$ and LR_{B} as given in (2.11) - (2.12); see Dagenais and Dufour (1991).

In order to establish the distribution of LR_{B}, it will be useful to first show that LR_{B} is invariant to data transformations of the form $\tilde{Y}_* = \tilde{Y}A$, where A is any non-singular fixed matrix of order n. Such transformations can also be viewed as affine transformations on Y with the following specific form:

$$Y_* = YA + \tilde{R}_{\mathrm{M}}\iota_n'(I_n - A). \tag{6.6}$$

Lemma 6.1 MULTIVARIATE SCALE INVARIANCE. *The LR statistics $LR(\gamma_0)$ and LR_{B} defined in* (2.11) *and* (2.12) *are invariant to replacing $\tilde{Y}$ by $\tilde{Y}_* = \tilde{Y}A$, where A is an arbitrary nonsingular $n \times n$ matrix.*

We can now establish the following theorem which characterizes the distribution of the test statistics $LR(\gamma_0)$ and LR_{B}, under both the null hypothesis $\mathcal{H}_{\mathrm{B}}$ and the corresponding unrestricted MLR alternative model.

Theorem 6.2 EXACT DISTRIBUTION OF BCAPM LR TESTS. *Under* (2.1) *and* (2.2), *the distributions of the statistics $LR(\gamma_0)$ and LR_{B} depend on the model parameters (B, K) only through $\bar{B} = (B - \Delta)K^{-1}$ where $\Delta = [0, \iota_n]'$, and*

$$LR(\gamma_0) = T\,\ln\left(|\hat{W}(\gamma_0)'\hat{W}(\gamma_0)|/|\hat{W}'\hat{W}|\right), \quad LR_{\mathrm{B}} = \inf\left\{LR(\gamma_0) : \gamma_0 \in \Gamma\right\}, \tag{6.7}$$

where $\hat{W}(\gamma_0) = \bar{M}(\gamma_0)(X\bar{B} + W) = \bar{M}(\gamma_0)\{\iota_T[a + \gamma_0(\beta - \iota_n)]'K^{-1} + W\}$, $\hat{W} = M(X)W$, and $\bar{M}(\gamma_0)$ is defined as in (4.1). *If, furthermore, the null hypothesis $\mathcal{H}_{\mathrm{B}}$ holds, then*

$$\hat{W}(\gamma_0) = (\gamma_0 - \gamma)\bar{M}(\gamma_0)\iota_T(\beta - \iota_n)'K^{-1} + \bar{M}(\gamma_0)W \tag{6.8}$$

and the distribution of LR_{B} depends on (B, K) only through γ and $(\beta - \iota_n)'K^{-1}$.

Even though the matrices B and K may involve up to $2n+ n^2$ different parameters [or $2n + n(n+1)/2$ parameters, if K is a triangular matrix], the latter theorem shows that the number of free parameters in the distributions of $LR(\gamma_0)$ and LR_{B} does not exceed $2n$, the number of parameters in $\bar{B}$. Indeed, when $\mathcal{H}_{\mathrm{B}}$ holds, the number of free parameters is at most $n+1$. Further, under $\mathcal{H}(\gamma_0)$, it follows from (6.8) that $\bar{B}$ completely disappears from the distribution, entailing Theorem **4.1**. However, Theorem **6.2** also provides the power function. The characterizations given by (6.7) and (6.8) are sufficiently explicit to allow simulating the distributions $LR(\gamma_0)$ and LR_{B} once $\bar{B}$ is specified.

Further information on the structure of the nuisance parameters can be drawn from considering the *singular value decomposition* of $\bar{B}$. Let r be the rank of $\bar{B}$. Since $\bar{B}$ is a $2 \times n$ matrix, we have $0 \leq r \leq 2$ and we can write:

$$\bar{B} = PDQ' , \quad D = [\bar{D}, 0] , \quad \bar{D} = \mathrm{diag}\left(\lambda_1^{1/2}, \lambda_2^{1/2}\right) , \tag{6.9}$$

where D is a $2 \times n$ matrix, λ_1 and λ_2 are the two largest eigenvalues of $\bar{B}'\bar{B}$ (where $\lambda_1 \geq \lambda_2 \geq 0$), $Q = [Q_1, Q_2]$ is an orthogonal $n \times n$ matrix whose columns are eigenvectors of $\bar{B}'\bar{B}$, Q_1 is a $2 \times r$ matrix which contains eigenvectors associated with the non-zero eigenvalues of $\bar{B}'\bar{B}$, $P = [P_1, P_2]$ is a 2×2 orthogonal matrix such that $P_1 = \bar{B}Q_1D_1^{-1}$ and D_1 is a diagonal matrix which contains the non-zero eigenvalues of $\bar{B}'\bar{B}$, setting $P = P_1$ and $D_1 = \bar{D}$ if $r = 2$, and $P = P_2$ if $r = 0$; see Harville (1997, Theorem 21.12.1).

Using Lemma **6.1** and Theorem **6.2**, $LR(\gamma_0)$ and LR_{B} may then be reexpressed as:

$$LR(\gamma_0) = T \, \ln\left(|\tilde{W}(\gamma_0)'\tilde{W}(\gamma_0)|/|\tilde{W}'\tilde{W}|\right) , \quad LR_{\mathrm{B}} = \inf\left\{LR(\gamma_0) : \gamma_0 \in \Gamma\right\} , \tag{6.10}$$

where $\tilde{W}(\gamma_0) = \hat{W}(\gamma_0)Q = \bar{M}(\gamma_0)(XPD + \bar{W})$, $PD = [P\bar{D},\, 0], \tilde{W} = \hat{W}Q = M(X)\bar{W}$ and $\bar{W} = WQ$. In view of the orthogonality of the P matrix, it is easy to see that the 2×2 matrix $P\bar{D}$ contains at most 3 free coefficients. Further, under $\mathcal{H}_{\mathrm{B}}$, the rank of $\bar{B}$ is at most one (so that $\lambda_2 = 0$) and $P\bar{D}$ involves at most one free coefficient.[13] When W follows a non-Gaussian distribution, the distributions of $LR(\gamma_0)$ and LR_{B} may be influenced by the parameters of $\bar{B}$ through the matrix Q in $\bar{W}$. However, under the Gaussian assumption $\mathcal{H}_W(\mathcal{D}_N)$ in (2.3), the rows of $\bar{W}$ are *i.i.d.* $\mathrm{N}(0,\, I_n)$, so that the statistics $LR(\gamma_0)$ and LR_{B} follow distributions which depend on (B, K) only through $P\bar{D}$. In particular, under $\mathcal{H}_{\mathrm{B}}$, this distribution involves only one nuisance parameter, in accordance with the result derived by Zhou (1991, Theorem 1) through a different method.

Even though the above observations can be useful – as we will see below – Theorem **6.2** leaves us with a distribution that depends on nuisance parameters under the null hypothesis $\mathcal{H}_{\mathrm{B}}$, possibly in a complex way. So bounds may be needed to test $\mathcal{H}_{\mathrm{B}}$. In the next section, we study the problem of testing $\mathcal{H}_{\mathrm{B}}$ despite the presence of nuisance parameters which may be difficult to evacuate.

[13]If $r = 1$, the eigenvector associated with λ_1 is uniquely defined once its first non-zero element is normalized to be positive (which is always possible).

7. Exact bound procedures for testing Black's CAPM

In this section, we propose procedures for testing $\mathcal{H}_{\mathrm{B}}$ despite the presence of nuisance parameters induced by the nonlinearity of $\mathcal{H}_{\mathrm{B}}$ and non-Gaussian error distributions. We study first global bounds based on tests of $\mathcal{H}(\gamma_0)$ discussed in Section 4, where we outline important differences between the Gaussian and non-Gaussian cases. Second, we describe a more general but computationally more expensive method based on the technique of MMC tests [Dufour (2006)] in order to obtain tighter bounds. Thirdly, for the case when the error distribution involves nuisance parameters which are difficult to eliminate, we propose variants of these approaches which involves building a confidence set for the nuisance parameters in the error distribution and may lead to tighter bounds.

7.1. Global bound induced by tests of $\mathcal{H}(\gamma_0)$

The results of Section 4 on testing $\gamma = \gamma_0$ can be used to derive a global bound on the distribution of the statistic LR_{B}. This is done in the following theorem.

Theorem 7.1 GLOBAL BOUND ON THE NULL DISTRIBUTION OF THE BCAPM TEST. *Under the assumptions* (2.1), (2.2) *and* (2.7), *the distribution of the statistic* LR_{B} *in* (2.11) *satisfies the following inequality for any given* $\nu \in \Omega_{\mathcal{D}}$:

$$\mathsf{P}\big[LR_{\mathrm{B}} \geq x\big] \leq \sup_{\gamma_0 \in \Gamma} \mathsf{P}\big[\,\overline{LR}(\gamma_0,\, W) \geq x\big]\,,\ \forall x\,, \tag{7.1}$$

where $\overline{LR}(\gamma_0,\, W)$ *is defined as in* (4.1). *Further, in the Gaussian case where* $W_1,\, \dots\, ,\, W_T \overset{i.i.d}{\sim} N$ $(0,\, I_n)$, *we have:*

$$\mathsf{P}\big[(T-1-n)\,(\Lambda_{\mathrm{B}}-1)\,/n \geq x\big] \leq \mathsf{P}[F(n,\, T-1-n) \geq x]\,,\ \forall x\,. \tag{7.2}$$

To relate this result to the bounds proposed by Shanken (1986) and Stewart (1997), first observe that (7.1) and (7.2) easily extend to the following multi-beta setups:

$$R_{it} = a_i + \sum_{j=1}^{s} \beta_{ij}\tilde{R}_{jt} + u_{it}\,, \quad i = 1,\, \dots\, ,\, n,\ t = 1,\, \dots\, ,\, T,$$

where $\tilde{R}_{jt}$, $j = 1,\, \dots\, ,s$ are returns on s benchmarks and the null hypothesis is

$$\mathcal{H}_{\mathrm{B}} : a_i = \gamma\Big(1 - \sum_{j=1}^{s} \beta_{ij}\Big)\,, \quad i = 1,\, \dots\, ,\, n. \tag{7.3}$$

In this case, the bounding distribution of the LR_{B} statistic obtains as in Theorems **7.1** where $X = [\iota_T,\, \tilde{R}_1,\, \dots\, ,\, \tilde{R}_s]$, $\tilde{R}_j = (\tilde{R}_{j1},\, \dots\, ,\, \tilde{R}_{jT})'$, $j = 1,\, \dots\, ,\, s$, and H is the k-dimensional row vector $(1,\, \gamma_0,\, \dots\, ,\, \gamma_0)$. In the Gaussian case, the probabilities $\mathsf{P}\big[\,\overline{LR}(\gamma_0,\, W) \geq x\big]$ do not depend on γ_0, and the bounding distribution under normality is $F(n,\, T-s-n)$. For this problem, Shanken

(1986) suggested the statistic

$$\tilde{Q} = \min_{\gamma} \left\{ \frac{T\big[\hat{a} - \gamma(\iota_n - \hat{\beta}\iota_s)\big]'\big[\big(T/(T-2)\big)\hat{\varSigma}\big]^{-1}\big[\hat{a} - \gamma(\iota_n - \hat{\beta}\iota_s)\big]}{1 + (\bar{R}_{\mathrm{M}} - \gamma\iota_s)'\hat{\varDelta}_{\mathrm{M}}^{-1}(\bar{R}_{\mathrm{M}} - \gamma\iota_s)} \right\} \tag{7.4}$$

where $\hat{a}$ is an n-dimensional vector which includes the (unconstrained) intercepts estimates, $\hat{\beta}$ is an $(n \times s)$ matrix whose rows includes the unconstrained OLS estimates of $(\beta_{i1}, \ldots, \beta_{is})$, $i = 1, \ldots, n$, $\bar{R}_{\mathrm{M}}$ and $\hat{\varDelta}_{\mathrm{M}}$ include respectively the time-series means and sample covariance matrix corresponding to the right-hand-side total portfolio returns. Further, the minimum in (7.4) occurs at the constrained ML estimator $\hat{\gamma}$ of γ, and LR_{B} is a transformation of $\hat{Q}$:

$$LR_{\mathrm{B}} = T \ln\left(1 + \frac{\hat{Q}}{T - s - 1}\right). \tag{7.5}$$

On assuming normal errors, the null distribution of $\hat{Q}$ may be bounded by the Hotelling distribution, *i.e.* $(T - s - n)\hat{Q}/[n(T - s - 1)]$ can be bounded by the $F(n, T - n - s)$ distribution. The latter obtains from Gibbons et al. (1989)'s joint test for the significance of the intercepts, where returns are expressed in excess of a known γ.

Independently, Stewart (1997) showed that the statistic $(T-s-n)\big[\big(|\hat{\varSigma}(\gamma)|/|\hat{\varSigma}|\big)-1\big]/n$ can be bounded, under normal errors, by the $F(n, T - n - s)$ distribution. This result may also be viewed as an instance of the bounding technique described in Dufour (1989). Now, from (2.11) and (7.5), we see easily that: (i) Shanken and Stewart's bounds are equivalent, and (ii) both results obtain from Theorem **7.1** in the special case of normal errors.

When disturbances are non-Gaussian, Theorem **7.1** entails that the bounding distribution can easily be simulated, as follows. Given a value of ν, generate N *i.i.d.* draws from the distribution of $W_1, \ldots, W_T$; then, for any given γ_0, these yield a vector $\overline{LR}_N(\gamma_0, \nu)$ of N simulated values of the test statistic $\overline{LR}(\gamma_0, W)$, as defined in (4.1). A MC p-value may then be computed from the rank of the observed statistic LR_{B} relative to the simulated values. Following the notation of Section 4, denote this MC p-value by

$$\hat{p}_N^U(\gamma_0, \nu) \equiv p_N[LR_{\mathrm{B}}^{(0)}|\overline{LR}_N(\gamma_0, \nu)] \tag{7.6}$$

where $LR_{\mathrm{B}}^{(0)}$ represents the value of the test statistic LR_{B} based on the observed data; we will call $\hat{p}_N^U(\gamma_0, \nu)$ the *bound MC* (BMC) p-value. Since $LR_{\mathrm{B}} \leq LR(\gamma_0)$, for any γ_0, we have $LR_{\mathrm{B}} \leq LR(\gamma)$ and $\hat{p}_N(\gamma, \nu) \leq \hat{p}_N^U(\gamma, \nu)$, hence by (4.6),

$$\mathsf{P}\big[\hat{p}_N^U(\gamma, \nu) \leq \alpha\big] \leq \mathsf{P}\big[\hat{p}_N(\gamma, \nu) \leq \alpha\big] = \alpha. \tag{7.7}$$

For any subsets $A \subseteq \varGamma$ and $E \subseteq \varOmega_{\mathcal{D}}$, set:

$$\hat{p}_N^U(\gamma_0, E) = \sup\{\hat{p}_N^U(\gamma_0, \nu_0) : \nu_0 \in E\},\ \hat{p}_N^U(A, \nu_0) = \sup\{\hat{p}_N^U(\gamma_0, \nu_0) : \gamma_0 \in A\},$$

$$\hat{p}_N^U(A, E) = \sup\{\hat{p}_N^U(\gamma_0, \nu_0) : \gamma_0 \in A,\ \nu_0 \in E\} \tag{7.8}$$

where, by convention, $\hat{p}_N^U(A, \cdot) = 0$ if A is empty, and $\hat{p}_N^U(\cdot, E) = 0$ if E is empty. In contrast with the Gaussian case where the bounding distribution is $F(n, T-n-p)$ for all γ_0, the distribution of $\hat{p}_N^U[\gamma_0, \nu]$ may depend on γ_0. Then a critical region that provably satisfies the level constraint can be obtained by maximizing $\hat{p}_N^U(\gamma_0, \nu)$ over the relevant nuisance parameters, as shown in the following theorem.

Theorem 7.2 GLOBAL SIMULATION-BASED BOUND ON THE NULL DISTRIBUTION OF THE BCAPM TEST STATISTIC. *Under* (2.1), (2.2) *and* (2.7),

$$\mathsf{P}\big[\hat{p}_N^U(\Gamma, \nu) \le \alpha\big] \le \alpha \tag{7.9}$$

where ν represents the true distributional shape of W, and

$$\mathsf{P}\big[\hat{p}_N^U(\Gamma, \Omega_{\mathcal{D}}) \le \alpha\big] \le \alpha\,. \tag{7.10}$$

In the above theorem, the critical region in (7.9) is applicable when the distributional shape parameter ν is specified (or known), while (7.10) holds even if ν is unknown. These bound tests are closely related to the CS-based test proposed in Section 5: the null hypothesis is rejected when the CS for γ is empty, *i.e.* if no value of γ_0 can deemed acceptable (at level α), either with ν specified or ν taken as a nuisance parameter. This may be seen on comparing (5.10) with (7.10). On recalling that $LR_{\mathrm{B}} = \inf\{LR(\gamma_0) : \gamma_0 \in \Gamma\}$, the latter also suggests a relatively easy way of showing that $C_\gamma(\alpha; \nu)$ or $C_\gamma(\alpha; \mathcal{D})$ is not empty, through the specific p-value $\hat{p}_N^U(\hat{\gamma}, \nu)$ obtained by taking $\gamma_0 = \hat{\gamma}$ in (7.6). We shall call $\hat{p}_N^U(\hat{\gamma}, \nu)$ the *QML-BMC* p-value.

Theorem 7.3 RELATION BETWEEN EFFICIENCY TESTS AND ZERO BETA CONFIDENCE SETS. *Under* (2.1), (2.2) *and* (2.7), *let $\hat{\gamma}$ be the QML estimator of γ in* (2.13). *Then,*

$$\hat{p}_N^U(\hat{\gamma}, \nu) > \alpha \Rightarrow \sup\{\hat{p}_N(\gamma_0, \nu) : \gamma_0 \in \Gamma\} > \alpha \Rightarrow C_\gamma(\alpha; \nu) \neq \emptyset\,,\ \forall \nu \in \Omega_{\mathcal{D}}, \tag{7.11}$$

$$\hat{p}_N^U(\hat{\gamma}, \Omega_{\mathcal{D}}) > \alpha \Rightarrow \sup\{\hat{p}_N(\gamma_0, \nu_0) : \gamma_0 \in \Gamma,\, \nu_0 \in \Omega_{\mathcal{D}}\} > \alpha \Rightarrow C_\gamma(\alpha; \mathcal{D}) \neq \emptyset\,, \tag{7.12}$$

where $C_\gamma(\alpha, \nu)$ and $C_\gamma(\alpha; \mathcal{D})$ are the sets defined in (5.9) *and* (5.10).

It is of interest to note here that, for the Gaussian special case where $\mathcal{H}_W(\mathcal{D}_N)$ holds, Zhou (1991) and Velu and Zhou (1999) proposed a bound based on a statistical result from Schott (1984, Theorem 3). Although potentially tighter than the bound proposed above, this result is applicable to statistics which can be written as ratios of independent Wishart variables and do not seem to extend easily to other classes of distributions. In the next section, we propose an approach which yields similarly tighter bounds for non-Gaussian distributions as well.

7.2. Maximized Monte Carlo procedures

We will now describe a tighter bound based on the MMC technique [Dufour (2006)]. To conform with our earlier notation for MC p-values, let us define the function $\overline{LR}_{\mathrm{B}}(\theta, W)\,, \quad \theta = \psi(B, K)$, which assigns to each value of model parameters B and K and the noise matrix W the following outcome: using θ and a draw from the distribution of W (which may depend on a nuisance parameter ν), generate a sample from model (2.1)-(2.7), and compute LR_{B} in (2.11) from this sample. The function $\theta = \psi(B, K)$ does represent the effective vector of nuisance parameters. For example, by Theorem **6.2**, the distribution of $\overline{LR}_{\mathrm{B}}$ under the null hypothesis $\mathcal{H}_{\mathrm{B}}$ depends on the model parameters (B, K) only through γ and $(\beta - \iota_n)'K^{-1}$, where γ is the solution of the equation $a = \gamma(\iota_n - \beta)$. Denote by Ω_{B} the set of admissible values for θ under $\mathcal{H}_{\mathrm{B}}$.

On applying $\overline{LR}_{\mathrm{B}}(\theta, W)$, we can get simulated values from the null distribution of LR_{B} for any value of θ. If N independent replications $W^{(1)}, \ldots, W^{(N)}$ of W are generated, we can then compute the corresponding vector of $\overline{LR}_{\mathrm{B}}$ statistics and the p-value function

$$\hat{p}_{\mathrm{B}N}(\theta, \nu) = p_N\left[LR_{\mathrm{B}}^{(0)} \middle| \overline{LR}_{\mathrm{B}N}(\theta, \nu)\right], \tag{7.13}$$

where $\overline{LR}_{\mathrm{B}N}(\theta, \nu) = \left[\overline{LR}_{\mathrm{B}}\big(\theta, W^{(1)}\big), \ldots, \overline{LR}_{\mathrm{B}}\big(\theta, W^{(N)}\big)\right]'$. For any given value of ν, the MMC p-value associated with $LR_{\mathrm{B}}^{(0)}$ is obtained by maximizing $\hat{p}_{\mathrm{B}N}(\theta, \nu)$ with respect to θ over the set of admissible values Ω_{B} under $\mathcal{H}_{\mathrm{B}}$:

$$\hat{p}_{\mathrm{B}N}^{M}(\Omega_{\mathrm{B}}, \nu) = \sup\left\{\hat{p}_{\mathrm{B}N}(\theta, \nu) : \theta \in \Omega_{\mathrm{B}}\right\}. \tag{7.14}$$

From the results in Dufour (2006), we have under $\mathcal{H}_{\mathrm{B}}$ and the error distribution associated with ν :

$$\mathsf{P}\left[\hat{p}_{\mathrm{B}N}^{M}(\Omega_{\mathrm{B}}, \nu) \leq \alpha\right] \leq \alpha\,; \tag{7.15}$$

in other words, $\hat{p}_{\mathrm{B}N}^{M}(\Omega_{\mathrm{B}}, \nu) \leq \alpha$ is a critical region with level α.

To allow for an unknown ν, we can maximize $\hat{p}_{\mathrm{B}N}(\theta, \nu)$ with respect to $\nu \in \Gamma_{\mathcal{D}}$. Set:

$$\hat{p}_{\mathrm{B}N}^{M}(\theta, \Omega_{\mathcal{D}}) = \sup\left\{\hat{p}_{\mathrm{B}N}(\theta, \nu) : \nu \in \Omega_{\mathcal{D}}\right\},\ \hat{p}_{\mathrm{B}N}^{M}(\Omega_{\mathrm{B}}, \Omega_{\mathcal{D}}) = \sup\left\{\hat{p}_{\mathrm{B}N}^{M}(\theta, \Omega_{\mathcal{D}}) : \theta \in \Omega_{\mathrm{B}}\right\}. \tag{7.16}$$

Then, under $\mathcal{H}_{\mathrm{B}}$, if

$$\mathsf{P}\left[\hat{p}_{\mathrm{B}N}^{M}(\Omega_{\mathrm{B}}, \Omega_{\mathcal{D}}) \leq \alpha\right] \leq \alpha\,. \tag{7.17}$$

Of course, the above p-values may be more costly to compute than those proposed in Section 7.1, but they are tighter by construction. Nevertheless, Theorem **7.3** guarantees that

$$\hat{p}_{N}^{U}(\Gamma, \nu) \leq \alpha \Rightarrow \hat{p}_{\mathrm{B}N}^{M}(\Omega_{\mathrm{B}}, \nu) \leq \alpha$$

for any given ν. So it may be useful to check the global bound for significance before turning to the MMC one. Furthermore, it is not always necessary to run the numerical p-value maximization underlying the MMC p-value to convergence: if $\hat{p}_{\mathrm{B}N}(\theta, \nu) > \alpha$ given any relevant θ (or ν), then a non-rejection is confirmed. So one may exit whenever a value larger than α is reached. We suggest to use the QML estimate $\hat{\theta}$ of θ as start-up value, because this provides *parametric bootstrap-type* [or a *local MC* (LMC)] p-values:

$$p_{N}^{b}(\nu) = \hat{p}_{\mathrm{B}N}(\hat{\theta}, \nu)\,,\ p_{N}^{b}(\Omega_{\mathcal{D}}) = \hat{p}_{\mathrm{B}N}(\hat{\theta}, \Omega_{\mathcal{D}})\,. \tag{7.18}$$

Then $p_N^b(\nu) > \alpha$ entails $\hat{p}_{\mathrm{B}N}^M(\Omega_{\mathrm{B}},\, \nu) > \alpha$, and $p_N^b(\,\Omega_{\mathcal{D}}) > \alpha$ entails $\hat{p}_{\mathrm{B}N}^M(\Omega_{\mathrm{B}},\, \Omega_{\mathcal{D}}) > \alpha$.

7.3. Bound two-stage confidence procedures

To deal with the fact that the distribution of W may involve an unknown parameter $\nu \in \Omega_{\mathcal{D}}$, we suggested above to maximize the relevant p-values over the nuisance parameter space $\Omega_{\mathcal{D}}$. Since this process may lead to power losses, we suggest to restrict the maximization over ν to a set which is empirically relevant. A way to do this consists in using an approach similar to the one proposed in Dufour (1990) and Dufour and Kiviet (1996). This method consists in two basic steps: (i) an exact CS with level $1-\alpha_1$ is built for ν, and (ii) the MC p-values (presented above) are maximized over all values of ν in the latter CS and are referred to the level α_2. This yields a test with level $\alpha = \alpha_1+\alpha_2$. In our empirical application, we used $\alpha/2$. Empirically, maximizing over a set estimator for ν (rather than over all the nuisance parameter set) ensures that the error distributions retained are consistent with our data. Recall that the null hypothesis imposes distributional constraints jointly with the restrictions on the regression coefficients. In view of this, retaining empirically relevant values for ν helps to interpret BCAPM test rejections.

Let $\mathcal{C}_\nu(\alpha_1) = \mathcal{C}_\nu(\alpha_1;\, Y)$ be a CS with level $1-\alpha_1$ for ν. Then, under the null hypothesis $\mathcal{H}(\gamma_0)$, we have

$$\mathsf{P}\big[\hat{p}_N^U[\gamma_0,\, \mathcal{C}_\nu(\alpha_1)] \le \alpha_2\big] \le \alpha_1+\alpha_2 \tag{7.19}$$

while, under $\mathcal{H}_{\mathrm{B}}$,

$$\mathsf{P}\big[\hat{p}_N^U[\Gamma,\, \mathcal{C}_\nu(\alpha_1)] \le \alpha_2\big] \le \alpha_1+\alpha_2\,, \tag{7.20}$$

$$\mathsf{P}\big[\hat{p}_{\mathrm{B}N}^M[\Omega_{\mathrm{B}},\, \mathcal{C}_\nu(\alpha_1)] \le \alpha_2\big] \le \alpha_1+\alpha_2\,. \tag{7.21}$$

Note also that

$$\hat{p}_{\mathrm{B}N}^M(\hat{\theta},\, \mathcal{C}_\nu(\alpha_1)) > \alpha_2 \tag{7.22}$$

is sufficient for $\hat{p}_{\mathrm{B}N}^M[\Omega_{\mathrm{B}},\, \mathcal{C}_\nu(\alpha_1)] \le \alpha_2$ *not to hold.*

Let us now present the set estimation method we propose to build CS's for ν. Again, this will be done by "inverting" a test (of level α_1) for the specification underlying (2.2) where $\nu = \nu_0$ for known ν_0, an approach which avoids the need to use regularity assumptions on ν. The test we invert is the three-stage MC GF test introduced in Dufour et al. (2003):

$$\mathrm{CSK} \;=\; 1-\min\left\{\hat{p}\left[\mathrm{ESK}(\nu_0)\right],\, \hat{p}\left[\mathrm{EKU}(\nu_0)\right]\right\} \tag{7.23}$$

$$\mathrm{ESK}(\nu_0) \;=\; \left|\mathrm{SK}-\overline{\mathrm{SK}}(\nu_0)\right|, \quad \mathrm{SK} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{i=1}^{T}\hat{d}_{ii}^3\,, \tag{7.24}$$

$$\mathrm{EKU}(\nu_0) \;=\; \left|\mathrm{KU}-\overline{\mathrm{KU}}(\nu_0)\right|, \quad \mathrm{KU} = \frac{1}{T}\sum_{t=1}^{T}\hat{d}_{tt}^2\,, \tag{7.25}$$

where $\hat{d}_{it}$ are the elements of the matrix $\hat{U}(\hat{U}'\hat{U}/T)^{-1}\hat{U}'$, $\overline{\mathrm{SK}}(\nu_0)$ and $\overline{\mathrm{KU}}(\nu_0)$ are simulation-based estimates of the expected SK and KU given (2.2) and $\hat{p}\left[\mathrm{ESK}(\nu_0)\right]$ and $\hat{p}\left[\mathrm{EKU}(\nu_0)\right]$ are p-values, obtained by MC methods under (2.2). The MC test technique is also applied to obtain a size correct

p-value for CSK. The CS for ν corresponds to the values of ν_0 which are not rejected at level α_1, using the latter p-value.[14]

8. Empirical analysis

From the empirical viewpoint, the literature on testing efficiency within the BCAPM setup remains scarce and has produced mixed results. While Gibbons (1982) rejected efficiency restrictions using an asymptotic critical value, Jobson and Korkie (1982), Shanken (1985, 1986) and Stewart (1997) found much higher p-values on applying finite-sample corrections (based respectively on a Bartlett correction and a small-sample Gaussian-based lower bound), which overall support the efficiency hypothesis.[15] This can be contrasted with the results of Zhou (1991) who provided empirical evidence on the efficiency of a portfolio using optimal upper and lower bounds computed with an eigenvalue-based test (under the Gaussian assumption). His evidence shows that efficiency is rejected in most subperiods. Velu and Zhou (1999) derived the exact distribution of the likelihood ratio test for the Gaussian multibeta pricing models in the absence of a risk-free rate and considered a GMM-based asymptotic test. These procedures were applied to size portfolios to test the efficiency of CRSP value-weighted and equally weighted indices and to the Fama and French (1993) multibeta model. Their results show that the efficiency of CRSP value-weighted and equally weighted indices cannot be rejected for all of the 10-year subperiods, except for the first one, while the single index efficiency is rejected in all subperiods. Furthermore, the tests are not favorable to the Fama and French (1993) multibeta model. Finally, Chou (2000) showed that the comparison of a Wald test, a GMM test, the CRST test of Shanken (1985) and a likelihood ratio test leads to the rejection of the Fama and French three-factor model. The results for the one factor model are mixed. It cannot be rejected for size portfolios but it is rejected for Fama and French portfolios. For industry portfolios, the one-factor model is rejected by the likelihood ratio test but not by the Wald and GMM tests. Given that no test outperforms the others no conclusions could be reached regarding the efficiency of CRSP index with respect to industry portfolios.

From a statistical viewpoint, there are at least three important observations one can draw from the theoretical and empirical literature on the BCAPM (and related models). First, due to the cross-sectional nature of the efficiency restrictions, there is no theoretical reason why coefficients should be constant over time: this feature must be taken into account in the empirical analysis, for example by considering relatively short subperiods.[16] Of course, this can easily lead to short samples. Second, the possibility of non-Gaussian – possibly heavy-tailed – error distributions must be modelled. Third, due to the multivariate nonlinear nature of the model, it is crucial that a good finite-sample distributional theory be supplied for the test statistics. In this section, we present an empirical analysis which takes these points into account.

[14] For additional discussion of the goodness-of-fit tests used for inference on ν, see Appendix D.

[15] In Shanken (1986), efficiency is accepted for two out of three six-year subperiods, using both an approximate and an upper bound on the p-value associated with likelihood ratio test.

[16] The importance of considering subperiods in empirical work has also been documented; see Black (1993), Fama and French (2004) and Beaulieu et al. (2006). See also the literature on conditional CAPM, *e.g.* MacKinlay and Richardson (1991), Jagannathan and Wang (1996) and Ferson and Harvey (1999).

We study the mean-variance efficiency tests of the market portfolio when the risk-free rate is not available [formally, tests of (2.7) in the context of the market model (2.1)] under error distributional assumptions (2.4)-(2.5). We use real monthly returns over the period going from January 1926 to December 1995, obtained from the University of Chicago's Center for Research in Security Prices (CRSP). As in Breeden, Gibbons and Litzenberger (1989), the data studied here involve 12 portfolios of New York Stock Exchange (NYSE) firms grouped by standard two-digit industrial classification (SIC). The sectors studied include: (1) petroleum; (2) finance and real estate; (3) consumer durables; (4) basic industries; (5) food and tobacco; (6) construction; (7) capital goods; (8) transportation; (9) utilities; (10) textile and trade; (11) services; (12) leisure; for details on the SIC codes, see Beaulieu et al. (2006). For each month the industry portfolios include the firms for which the return, price per common share and number of shares outstanding are recorded by CRSP. Furthermore, portfolios are value-weighted in each month. We measure the market return by the value-weighted NYSE returns, also available from CRSP. The real risk-free rate is proxied by the one month Treasury bill rate also from CRSP net of inflation. All MC tests were applied with 999 replications. All maximized MC p-values are obtained using the simulated annealing algorithm [Goffe, Ferrier and Rogers (1994)].[17] As is usual in this literature, given documented temporal instabilities, we estimate and test the model over 5 year subperiods.

Our BCAPM test results are summarized in Table 2. Non-Gaussian p-values are the largest MC p-values over the error distribution parameters [respectively: κ and (π, ω)] within the specified CS (with level 95%); the latter are reported in Table 2. A few useful guidelines are worth emphasizing: (i) given a 5% level test, the benchmark is .05 for p_∞ and $p_{\mathcal{N}}$, while for the Student t and normal mixture results the benchmark is .025; (ii) non-rejections using the LMC MC p-values [reported in columns (3) and (6), and (9)] are conclusive (though rejections are not); (iii) rejections based on the F-based conservative bound reported in column (5) are conclusive under normality; (iv) non-rejections based on the QML BMC p-value in the non-Gaussian case [reported in columns (8) and (11)] signal that the CS for γ is not empty; however, since the MMC p-value is based on the tightest bound, this evidence does not necessarily imply acceptance of $\mathcal{H}_{\mathrm{B}}$.

The empirical results presented here show that the asymptotic test and the Gaussian based bound test yield the same decision at 5% level, although the former p-values are much lower. Furthermore, the non-normal p-values exceed the Gaussian-based p-value, enough to change the test decision. It is relatively "easy" to reject the testable implications under normality. For instance, at the 5% significance level, we find seven rejections of the null hypothesis for the asymptotic $\chi^2(11)$ test, seven for the MC p-values under normality, and five (relying on the MMC p-value) under the Student-t and normal mixture distributions. We see that, focusing on t and normal mixture distributions with parameters not rejected by proper GF tests, mean-variance efficiency test results can change relative to the available F-based bound. The power advantages of the MMC procedure are illustrated by the results of the 1966-70 subperiod where the QML p-value exceeds 2.5% for the Student-t distribution, whereas the MMC p-value signals a rejection.

The CS's for distributional parameters are reported in Table 3. In the mixtures case, the confi-

[17]This search algorithm does not require a gradient function and sweeps the parameter space according to a randomized scheme. To escape local minima, a downhill step is always accepted while an uphill step may be accepted; the direction of all moves is determined by probabilistic criteria.

Table 2. Asymptotic and exact tests of BCAPM

Sample	(1) LR	(2) p_∞	(3) Normal LMC	(4) Normal MMC	(5) Normal BND	(6) Student t LMC	(7) Student t MMC	(8) Student t BND	(9) Normal mixture LMC	(10) Normal mixture MMC	(11) Normal mixture BND
1927-30	16.10	.137	.269	.308	.366	.272	.316	.360	.279	.313	.381
1931-35	14.09	.228	.344	.381	.432	.359	.399	.468	.342	.387	.452
1936-40	15.36	.167	.257	.284	.345	.282	.308	.372	.265	.302	.357
1941-45	18.62	.068	.148	.163	.203	.147	.169	.210	.150	.165	.211
1946-50	32.69	**.001**	.005	**.006**	**.007**	.007	**.007**	.010	.007	**.007**	.008
1951-55	37.04	**.000**	.003	**.004**	**.004**	.003	**.005**	.005	.003	**.003**	.003
1956-60	26.10	**.006**	.027	**.031**	**.042**	.030	.040	.052	.028	.035	.045
1961-65	29.21	**.002**	.011	**.016**	**.020**	.013	**.017**	.023	.014	**.021**	.024
1966-70	27.45	**.004**	.016	**.018**	**.026**	.020	**.025**	.032	.018	**.023**	.028
1971-75	16.81	.113	.213	.224	.292	.217	.248	.300	.206	.238	.292
1976-80	25.76	**.007**	.027	**.031**	**.040**	.026	.035	.039	.026	.034	.042
1981-85	14.98	.183	.316	.335	.387	.323	.399	.405	.318	.339	.406
1986-90	35.41	**.000**	.003	**.004**	**.004**	.004	**.005**	.005	.004	**.004**	.005
1991-95	16.41	.127	.219	.253	.310	.226	.263	.325	.226	.261	.319

Note – Column (1) presents the quasi-LR statistic defined in (2.11) to test $\mathcal{H}_{\mathrm{B}}$ [see (2.7)]; the remaining columns report the associated p-values using, respectively, the asymptotic $\chi^2(n-1)$ distribution, the Gaussian based MC p-values [columns (3)–(5)] and the MC p-values imposing multivariate $t(\kappa)$ errors [columns (6)–(8)] and mixture-of-normals $(\pi,\, \omega)$ errors [columns (9)–(11)]. Numbers in columns (6)-(11) are the largest MC p-values over the error distribution parameters [respectively: κ and $(\pi,\, \omega)$] within the specified CS's; the latter are reported in Table 3. In particular, the LMC p-values [columns (3), (6) and (9)] correspond to $\hat{p}^M_{\mathrm{B}N}(\hat{\theta},\, \mathcal{C}_\nu(\alpha_1))$ in (7.22). The MMC p-values [columns (4), (7) and (10)] correspond to $\hat{p}^M_{\mathrm{B}N}[\Omega_{\mathrm{B}},\, \mathcal{C}_\nu(\alpha_1)]$ in (7.21). Columns (5) report the Gaussian bound p-value corresponds based on (7.2); columns (8) and (11) report the QML-BMC p-values $\hat{p}^U_N[\hat{\gamma},\, \mathcal{C}_\nu(\alpha_1)]$ in (7.22). Returns for the month of January and October 1987 are excluded from the data set. Given a 5% level test, the benchmark is .05 for p_∞ and $p_{\mathcal{N}}$, while for the Student-t and normal mixture results the benchmark is .025. p-values which lead to significant tests with this benchmark are marked in bold.

Table 3. Confidence sets for intervening parameters

	Mixture $(\pi,\, \omega)$, confidence set for ω					$t(\kappa)$
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	$\pi = 0.1$	$\pi = 0.2$	$\pi = 0.3$	$\pi = 0.4$	$\pi = 0.5$	κ
1927-30	$\geq$1.8	1.6-2.8	1.6-2.5	1.6-2.5	1.6-2.6	3-12
1931-35	2.1-10.0	1.9-3.0	1.9-2.7	1.9-2.7	2.1-3.0	3-8
1966-40	1.5-3.5	1.5-2.3	1.4-2.1	1.4-2.0	1.4-2.1	4-25
1941-45	1.3-3.5	1.3-2.1	1.3-1.9	1.3-1.8	1.3-1.9	≥ 5
1946-50	1.4-3.5	1.3-2.2	1.3-2.0	1.3-1.9	1.3-1.9	5-37
1951-55	1.4-3.5	1.4-2.2	1.3-2.0	1.3-1.9	1.3-2.0	5-34
1956-60	1.3-2.8	1.2-2.0	1.2-1.9	1.2-1.8	1.2-1.8	≥ 5
1961-65	1.0-2.2	1.0-1.6	1.0-1.5	1.0-1.5	1.0-1.5	≥ 7
1966-70	1.3-3.0	1.3-2.0	1.3-1.9	1.3-1.8	1.2-1.9	≥ 5
1971-75	1.5-3.5	1.5-2.2	1.4-2.0	1.4-1.9	1.4-2.0	4-24
1976-80	1.6-4.0	1.5-2.5	1.5-2.2	1.5-2.2	1.5-2.3	4-19
1981-85	1.4-3.5	1.4-2.1	1.3-2.0	1.3-1.9	1.4-2.0	5-33
1986-90	1.1-3.0	1.1-2.0	1.1-1.8	1.0-1.7	1.1-1.8	≥ 5
1991-95	1.0-1.9	1.0-1.5	1.0-1.4	1.0-1.3	1.0-1.3	≥ 19

Note – Numbers in columns (1)-(5) represent a CS for the parameters $(\pi,\, \omega)$ [respectively, the probability of mixing and the ratio of scales] of the multivariate mixtures-of-normal error distribution. Column (6) presents the CS for κ, the degrees-of-freedom parameter of the multivariate Student-t error distribution. See Section 7 for details on the construction of these CS's: the values of $(\pi,\, \omega)$ or κ (respectively) in this set are not rejected by the CSK test (7.23) [see Dufour et al. (2003)] under multivariate mixtures or Student-t errors (respectively). Note that the maximum of the p-value occurs in the closed interval for ω. Returns for the month of January and October 1987 are excluded from the data set.

dence region is summarized as follows for presentation ease: we give the CS for ω corresponding to five different values of π namely 0.1, 0.2, 0.3, 0.4 and 0.5.

In Table 4, we present the average real market return as well as the average real risk-free rate over each subperiod, the QML estimate of γ and 95% CS's for this parameter, using, respectively, the asymptotic standard errors (3.1) [column (3)], the MC Gaussian based Fieller method [column (4)], and the MMC Fieller method imposing $t(\kappa)$ errors and normal mixture $(\pi,\, \omega)$ errors [columns (5)–(6)]. Note that some values of $\hat{\gamma}$ are high. Nonetheless, comparing the average real market return for those subperiods with our estimate of γ reveal that these high occurrences of γ are consistent with subperiods during which the estimated zero-beta rate is higher than the market portfolio return. This is an illustration of finding, ex post, a linear relationship between risk and return with a negative slope. Furthermore, rerunning our analysis using 10-year subperiods leads to γ estimates below the benchmark average return. The values of γ in the Fieller-based set are not rejected by the test defined in Theorem **4.1** to test $\mathcal{H}\,(\gamma_0)$. Rejection decision are based on the largest MC p-values over all κ and $(\pi,\, \omega)$ respectively; we did not restrict maximization to the CS for these parameters here. As expected in view of the $\mathcal{H}_{\mathrm{B}}$ test results, the exact CS's are empty for several subperiods.

Table 4. Point and set estimates for the zero–beta portfolio rate

			Estimate	95% Confidence set			
			QML	Wald-type	MMC based Fieller-type		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	$\bar{R}_M$	$\bar{r}_f$	$\hat{\gamma}$	Asymptotic	Normal errors	Student t errors	Mixture errors
1927-30	.0045	.0045	.0047	$[-.0037\,,\ .0130]$	$[-.0133\,,\ .0227]$	$[-.0143\,,\ .0229]$	$[-.0141\,,\ .0227]$
1931-35	.0103	.0025	-.0130	$[-.0301\,,\ .0039]$	$[-.0509\,,\ .0225]$	$[-.0520\,,\ .0225]$	$[-.0157\,,\ .0227]$
1926-40	.0031	-.0006	-.0069	$[-.0192\,,\ .0055]$	$[-.0341\,,\ .0187]$	$[-.0350\,,\ .0190]$	$[-.0349\,,\ .0817]$
1941-45	.0097	-.0042	.0117	$[.0037\,,\ .0198]$	$[-.0045\,,\ .0275]$	$[-.0048\,,\ .0287]$	$[-.0045\,,\ .0283]$
1946-50	.0021	-.0051	-.0219	$[-.0189\,,\ -.0070]$	$\emptyset$	$\emptyset$	$\emptyset$
1951-55	.0145	.0001	.0024	$[-.0015\,,\ .0064]$	$\emptyset$	$\emptyset$	$\emptyset$
1956-60	.0086	.0002	.0156	$[.0109\,,\ .0202]$	$\emptyset$	$[.0149\,,\ .0161]$	$\emptyset$
1961-65	.0080	.0014	.0571	$[.0398\,,\ .0744]$	$\emptyset$	$\emptyset$	$\emptyset$
1966-70	.0008	.0004	.0169	$[.0096\,,\ .0242]$	$\emptyset$	$\emptyset$	$\emptyset$
1971-75	-.0061	-.0010	.0150	$[.0030\,,\ .0270]$	$[-.0069\,,\ .0454]$	$[-.0081\,,\ .0488]$	$[-.0069\,,\ .0531]$
1976-80	.0056	-.0012	-.0096	$[-.0169\,,\ -.0024]$	$\emptyset$	$\emptyset$	$\emptyset$
1981-85	.0081	.0037	.0197	$[.0125\,,\ .0268]$	$[.0059\,,\ .0371]$	$[.0051\,,\ .0376]$	$[.0051\,,\ .0387]$
1986-90	.0088	.0020	.0053	$[-.0024\,,\ .0131]$	$\emptyset$	$\emptyset$	$\emptyset$
1991-95	.0104	.0011	.0010	$[-.0130\,,\ .0062]$	$[-.0285\,,\ .0147]$	$[-.0303\,,\ .0154]$	$[-.0325\,,\ .0147]$

Note – Column (1) shows the average real market portfolio return for each subperiod, column (2) the real average risk-free rate for each subperiod; column (3) presents $\hat{\gamma}$, the QML estimate of γ; the remaining columns report 95% CS's for this parameter, using, respectively, the asymptotic standard errors (3.1) [column (4)], the MC Gaussian based Fieller method [column (5)], and the MMC Fieller method imposing multivariate $t(\kappa)$ errors and mixture-of-normals $(\pi,\,\omega)$ errors [columns (6)–(7)]. See Section 5 for details on the construction of these CS's: the values of γ in this set are not rejected by the QLR test $LR(\gamma_0)$ defined in Theorem **4.1** to test the null hypotheses $\mathcal{H}(\gamma_0)$ for known γ_0. Values in columns (6)-(7) are based on the largest MC p-values over all relevant intervening parameters [respectively: κ and $(\pi,\,\omega)$]. Returns for the months of January and October 1987 are excluded from the data set.

The usefulness of the asymptotic confidence intervals is obviously questionable here, since our tests have led to the conclusion that no value of γ as defined by the model is consistent with our data. Other results which deserve notice are the empty sets for 1956-60 subperiod. These sets correspond to the case where the efficiency bound test is significant (at 5%).[18]

To further illustrate the differences between the asymptotic and our CS, we next check whether the average real risk-free rate is contained in these CS's (respectively). We clearly see that for many subperiods like 1966-70 the evidence is consistent between the asymptotic and MC Fieller-type confidence intervals. There are nonetheless cases where the set estimates do not lead to the same decision. For instance, in the 1941-45 and 1971-75 subperiods, the average risk-free rate is not included in the asymptotic confidence interval while it is covered by our MMC based Fieller type confidence intervals. These are cases where, using the asymptotic confidence interval, the hypothesis $\gamma = r_f$ is rejected, whereas exact confidence intervals indicate that it should not be rejected. Conversely, in 1986-90, the asymptotic confidence interval includes the average risk-free rate, whereas our CS's are empty (which provides exactly non-spurious evidence against our hypothesis).

We next perform multivariate misspecification tests: exact tests of multivariate normality [Dufour et al. (2003)], tests for GARCH effects [generalized Engle and Lee-King tests, based on extensions of the univariate tests proposed by Engle (1982) and Lee and King (1993))] and multivariate variance ratio tests [based on the univariate tests of Lo and MacKinlay (1988), Lo and MacKinlay (1989)] proposed in Dufour, Khalaf and Beaulieu (2005) and Beaulieu et al. (2006). The normality tests [see (7.23)] allow one to evaluate whether observed residuals exhibit non-normal behavior through excess skewness (ESK) and excess kurtosis (EKU), as well as jointly using our proposed combined test (CSK). p-values are reported in columns (1)-(3) of Table 5. The GARCH and variance ratio tests can be summarized as follows. We first obtain a standardized residuals matrix by multiplying the OLS residuals with a Cholesky type decomposition of $\hat{\Sigma}$. Using these modified residuals, we calculate individual-equation GARCH and variance ratio statistics then combine across-equations applying the same minimum p-value criterion used for the multi-normality test.[19]The MC test procedure here allows one to avoid using the Bonferroni-Boole bound [as done, for example, by Shanken (1990)], hence eliminating a potential power loss (without loosing the exactness of the test). To account for an unknown ν (given non-normal errors), we use the same approach as in the BCAPM test: we obtain the largest MC p-value overall relevant ν; the same CS for ν we applied for BCAPM is used. We consider 12 lags for all tests. The results of these tests are reported in columns 4-12 of Table 5.

For most subperiods, multivariate normality is rejected. These results confirm the results of Richardson and Smith (1993) who provide (asymptotic) evidence against multivariate normality in monthly returns. Of course, this evidence further justifies our approach to test the mean-variance efficiency of the market portfolio in BCAPM context under non-Gaussian errors. The results of the GARCH and variance ratio tests show very few rejections of the null hypothesis both at the 1%

[18]This can be checked by referring to Table 2: although the reported maximal p-values $[Q_U]$ in this table are performed over the confidence set for κ and (π, ω), we have checked that the global maximal p-value leads to the same decision here. Actually, it is a fact that the global maxima were observed to be very close to those obtained over the confidence set for the BMC test. It is not so for the diagnostic tests where the p-values were more variable.

[19]For further details on these tests, see Appendix E.

Table 5. Multivariate diagnostics

Error distribution	Normality test			Multivariate GARCH and variance ratio test								
				Gaussian			Student t			Normal mixture		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	SK	KU	CSK	$\widetilde{E}$	$\widetilde{LK}$	$\widetilde{VR}$	$\widetilde{E}$	$\widetilde{LK}$	$\widetilde{VR}$	$\widetilde{E}$	$\widetilde{LK}$	$\widetilde{VR}$
1927-30	**.001**	**.001**	**.001**	**.001**	.276	**.004**	**.015**	.228	**.004**	.159	.254	**.004**
1931-35	**.001**	**.001**	**.001**	**.022**	.697	.070	.082	.613	.065	.167	.624	.083
1926-40	**.001**	**.001**	**.001**	.064	.637	.853	.107	.616	.864	.136	.620	.867
1941-45	**.004**	**.002**	**.004**	.834	.984	.174	.847	.986	.184	.837	.983	.182
1946-50	**.002**	**.002**	**.002**	.166	.831	.096	.209	.810	.117	.245	.817	.112
1951-55	**.002**	**.001**	**.001**	.141	.344	.128	.179	.336	.146	.200	.320	.145
1956-60	**.027**	**.003**	**.005**	.986	.628	.092	.994	.626	.093	.991	.626	.094
1961-65	.613	.476	.625	.343	.174	.578	.382	.169	.586	.356	.180	.590
1966-70	**.012**	**.002**	**.005**	**.026**	.293	.826	.043	.302	.848	.073	.302	.850
1971-75	**.001**	**.002**	**.001**	.295	.215	.208	.337	.204	.219	.336	.202	.212
1976-80	**.001**	**.002**	**.001**	**.003**	**.011**	.167	**.011**	**.013**	.185	.042	**.010**	.184
1981-85	**.001**	**.002**	**.001**	**.035**	.100	.217	.060	.104	.225	.091	.094	.229
1986-90	**.027**	**.018**	**.025**	**.026**	.455	.311	.072	.450	.331	.081	.457	.323
1991-95	.167	.299	.226	.797	.244	.088	.823	.238	.093	.814	.250	.095

Note – Numbers in columns (1)-(3) represent p-values for multinormality tests: numbers in (1)-(2) pertain to the null hypothesis of respectively no excess skewness and no excess kurtosis in the residuals of each subperiod. The p-values in column (3) correspond to the combined statistic CSK designed for joint tests of the presence of skewness and kurtosis, for multivariate normal errors. p-values are MC pivotal statistics based. Numbers shown in columns (4)–(12) are p-values associated with the combined tests $\widetilde{E}$, $\widetilde{LK}$, and $\widetilde{VR}$, described in Dufour et al. (2003). $\widetilde{E}$ and $\widetilde{LK}$ are multivariate versions of Engle's and Lee and King's GARCH tests; $\widetilde{VR}$ is a multivariate version of Lo and MacKinlay's variance ratio tests. All p-values in columns (4)-(12) are the largest MC p-values over all intervening parameters [respectively: κ and $(\pi,\,\omega)$] within the specified CS's. The relevant 2.5% CS's for the nuisance parameters are reported in Table 3. Returns for the month of January and October 1987 are excluded from the data set.

and 5% level of significance. This implies that, in our statistical framework, *i.i.d.* errors provide an acceptable working assumption.

9. Conclusion

In this paper, we have shown that, in both Gaussian or non-Gaussian contexts, exact tests may be developed and used to assess the mean-variance efficiency of the market portfolio when the risk-free rate is not observable. After observing that Black's CAPM raises identification difficulties similar to the ones studied in the recent literature on weak identification, we proposed econometric methods which are robust to identification problems and allow for possibly heavy-tailed return distributions, such as the Student-t or normal mixtures. The approach described combines bound and MC test methods, and guarantees significance level control. Two earlier Gaussian based bound tests [Shanken (1986), Stewart (1997)] can also be viewed as special cases of our procedure.

As a complement to efficiency tests, we proposed exact Fieller-type CS's for zero-beta rate γ, which differ in a fundamental way from standard Wald-type confidence intervals. While Wald-type intervals can be highly unreliable and lead to substantially different inference concerning γ, *e.g.*, on whether γ is equal to the average real risk-free rate – a feature entailed by both econometric theory and simulation results – the Fieller-type intervals are provably valid in finite samples irrespective of the presence of identification problems. Interestingly, the latter are associated with a specific BCAPM test, which entails an empty CS for γ if the BCAPM efficiency restriction is not compatible with the data. Indeed, efficiency implies the existence of a scalar γ which equates the ratio of each $alpha$ coefficient to the corresponding $(1 - beta)$ coefficient. If the CS is empty, no value of γ is consistent with the data, which means that the BCAPM should be rejected. Apart from their usual interpretations, the proposed CS's may be used as an alternative efficiency test or to assess parameter stability over time. Of course, this is quite relevant for financial applications which use the BCAPM and require reliable estimates of γ.

The methods proposed were applied to monthly returns on 12 portfolios of the New York Stock Exchange (NYSE) firms over the period 1925-1995 divided in five-year subperiods. The results obtained include the following findings: (i) multivariate normality is rejected in most subperiods, so that results based on the Gaussian distributional assumption are highly questionable; (ii) multivariate residual checks reveal no significant departures from the *i.i.d.* assumption; (iii) the exact CS's for the zero-beta rate are very different from the asymptotic ones; (iv) mean-variance efficiency of the benchmark portfolio cannot be rejected over most subperiods (10 among 14 subperiods) once we allow for the possibility of non-normal errors and use finite-sample p-values; by contrast, Gaussian-based tests lead to rejections over 7 subperiods. Although there is evidence that further improvement of the model is warranted – for example through better modelling of the disturbance distribution and the use of conditioning information – these findings suggest that a statistically satisfactory specification of the Black's CAPM may be developed.

Appendix

A. Exact tests of uniform linear hypotheses

In this Appendix, we summarize general exact results from Dufour and Khalaf (2002) regarding hypotheses tests for any market model which may be cast in terms of the model (2.9) under assumption (2.2). Two classes of hypotheses are considered: (i) a general possibly nonlinear hypothesis

$$\mathcal{H}_{NL} : \check{R}\,[vec(B)] \in \Omega \tag{A.1}$$

where $\check{R}$ is a $\check{r} \times (nk)$ matrix of rank $\check{r}$, and Ω is a non-empty subset of $\Re^{\check{r}}$, and (ii) a *uniform linear* [UL] special case

$$\mathcal{H}_{UL} : HB = D\,, \tag{A.2}$$

where H is an $h \times k$ matrix of rank h. These hypotheses are relevant for BCAPM tests since the CAPM hypothesis H_{BCAPM} in (2.7) can be rewritten as:

$$(1,\,\gamma)B = \gamma\iota_n' \quad \text{for some (unknown) } \gamma\,, \tag{A.3}$$

which means it is of the form (A.1). Further, the special case (2.8) which corresponds to

$$(1,\,\gamma_0)B = \gamma_0\iota_n' \tag{A.4}$$

where γ_0 is known, takes the UL form (A.2) with $h = 1$. In this context, the (Gaussian) QLR criteria associated with H_{NL} and H_{UL} are, respectively,

$$LR_{NL} = T\ln(\Lambda_{NL})\,, \quad \Lambda_{NL} = |\hat{\Sigma}_{NL}|/|\hat{\Sigma}|\,, \tag{A.5}$$
$$LR_{UL} = T\ln(\Lambda_{UL})\,, \quad \Lambda_{UL} = |\hat{\Sigma}_{UL}|/|\hat{\Sigma}|\,, \tag{A.6}$$

where $\hat{\Sigma}_{NL}$ and $\hat{\Sigma}_{UL}$ are the QML estimators of Σ imposing (A.1) and (A.2) respectively, and $\hat{\Sigma}$ is the unconstrained QMLE.

Theorem A.1 DISTRIBUTION OF THE QUASI-LR UNIFORM LINEAR HYPOTHESIS TEST STATISTIC. *Under* (2.2)*,* (2.9) *and* (A.2)*, the Gaussian QML statistic* (A.6) *is distributed like*

$$\overline{LR} = T\ln\left(\left|W'M_{UL}W\right| / \left|W'MW\right|\right) \tag{A.7}$$

where $M_{UL} = M + X(X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H(X'X)^{-1}X'$, $M = I - X(X'X)^{-1}X$ *and* $W = [W_1,\ \dots\ ,\ W_n]$ *is defined by* (2.2).

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

$$\frac{(T-(k-1)-n)}{n}(\Lambda_{UL} - 1) \sim F(n,\,T-(k-1)-n)\,. \tag{A.8}$$

In this case, the Hotelling's T^2 criterion is a monotonic function of Λ_{UL}.

Theorem A.2 GENERAL SIMULATION BASED BOUND ON THE NULL DISTRIBUTION OF THE QUASI-LR TEST STATISTIC. *Under* (2.9), (2.2) *and* (A.1)*, the null distribution of the Gaussian QML statistic* (A.5) *may be bounded as follows:*

$$\mathsf{P}[LR_{NL} \geq x] \leq \mathsf{P}\left[T \ln\left(\frac{|W' M_{UL} W|}{|W' M W|}\right) \geq x\right] \, , \; \forall x \, , \tag{A.9}$$

where $M = I - X(X'X)^{-1}X'$, $M_{UL} = M + X(X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}H(X'X)^{-1}X'$, H *is an* $h \times k$ *matrix of rank* h *which satisfies*

$$\mathcal{H}_{ULBOUND} : \{HB = D\} \subseteq \mathcal{H}_{NL} \, , \tag{A.10}$$

and $W = [W_1, \ldots , W_p]$ *is defined by* (2.2).

This result obtains using the following arguments. Let $LR_{ULBOUND}$ refer to the QLR statistic to test (A.10), which is UL [*i.e.* of the form (A.2)]. By construction,

$$\mathcal{H}_{ULBOUND} \subseteq \mathcal{H}_{NL} \Rightarrow LR_{NL} \leq LR_{ULBOUND}. \tag{A.11}$$

This implies that $P[LR_{NL} \geq x] \leq P[LR_{ULBOUND} \geq x]$, for all x, and establishes the desired bound using Theorem **A.1**.

B. Proofs

PROOF OF THEOREM **4.1** Under (2.2) and $\mathcal{H}(\gamma_0)$, we have:

$$T\hat{\Sigma} = \hat{U}'\hat{U} = K'W'MWK \, , \quad T\hat{\Sigma}(\gamma_0) = K'W'\bar{M}(\gamma_0)WK \, .$$

Then, under $\mathrm{H}(\gamma_0)$,

$$\Lambda(\gamma_0) = \frac{|\hat{\Sigma}(\gamma_0)|}{|\hat{\Sigma}|} = \frac{|K'W'\bar{M}(\gamma_0)WK|}{|K'W'MWK|} = \frac{|K'| \, |W'\bar{M}(\gamma_0)W| \, |K|}{|K'| \, |W'MW| \, |K|} = \frac{|W'\bar{M}(\gamma_0)W|}{|W'MW|} \, ,$$

hence $P[LR(\gamma_0) \geq x] = P[T \ln(|W'\bar{M}(\gamma_0)W| \, / \, |W'MW|) \geq x]$, $\forall x$. □

PROOF OF LEMMA **6.1** The Gaussian log-likelihood function for model (6.2) is

$$\ln[\tilde{L}(\tilde{Y}, C \, , \Sigma)] = -\frac{T}{2}[n(2\pi) + \ln(|\Sigma|)] - \frac{1}{2}\mathrm{tr}[\Sigma^{-1}(\tilde{Y} - XC)'(\tilde{Y} - XC)] = \ln[L(Y, B \, , \Sigma)] \, .$$

Setting $\tilde{\Sigma}(C) \equiv \frac{1}{T}(\tilde{Y} - XC)'(\tilde{Y} - XC)$, for any given value of C, $\ln[\tilde{L}(\tilde{Y}, C \, , \Sigma)]$ is maximized by taking $\Sigma = \tilde{\Sigma}(C)$ yielding the concentrated log-likelihood

$$\ln[\tilde{L}(\tilde{Y},\, C\,,\, \Sigma)_c = -\frac{nT}{2}[(2\pi) + 1] - \frac{T}{2}\ln(|\tilde{\Sigma}(C)|)\,. \tag{B.1}$$

The Gaussian ML estimator of C thus minimizes $|\tilde{\Sigma}(C)|$ with respect to C. Let us denote by $\hat{C}(Y)$ the unrestricted ML estimator of C so obtained, and by $\hat{C}(Y;\, \gamma_0)$ and $\hat{C}_{\mathrm{B}}(Y)$ the restricted estimators subject to $\tilde{\mathcal{H}}(\gamma_0)$ and $\tilde{\mathcal{H}}_{\mathrm{B}}$ respectively.

Suppose that $\tilde{Y}$ is replaced by $\tilde{Y}_* = \tilde{Y}A$ where A is a non-singular $n \times n$ matrix. We need to show that $LR_*(\gamma_0) = LR(\gamma_0)$ and $LR_{\mathrm{B}*} = LR_{\mathrm{B}}$, where $LR_*(\gamma_0)$ and $LR_{\mathrm{B}*}$ represent the corresponding test statistics based on the transformed data. Following this transformation, $|\tilde{\Sigma}(C)|$ becomes:

$$\begin{aligned} |\tilde{\Sigma}_*(C_*)| &= \big|\frac{1}{T}(\tilde{Y}_* - XC_*)'(\tilde{Y}_* - XC_*)\big| = \big|\frac{1}{T}A'(\tilde{Y} - XC_*A^{-1})'(\tilde{Y} - XC_*A^{-1})A\big| \\ &= |A'A|\,\big|\frac{1}{T}(\tilde{Y} - XC)'(\tilde{Y} - XC)\big| = |A'A|\,|\tilde{\Sigma}(C)| \end{aligned} \tag{B.2}$$

where $C = C_*A^{-1}$. Then $|\tilde{\Sigma}(C_*)|$ is minimized by $\hat{C}_*(Y_*) = \hat{C}(Y)A$ and $|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)| = |A'A||\tilde{\Sigma}\big(\hat{C}(Y)\big)$. Similarly, on observing that

$$H(\gamma_0)C = 0 \Longleftrightarrow H(\gamma_0)CA = 0 \Longleftrightarrow H(\gamma_0)C_* = 0 \tag{B.3}$$

for any γ_0, it follows that the restricted estimators of C under $\tilde{\mathcal{H}}(\gamma_0)$ and $\tilde{\mathcal{H}}_{\mathrm{B}}$ are transformed in the same way: $\hat{C}_*(Y_*;\, \gamma_0) = \hat{C}(Y;\, \gamma_0)A$ and $\hat{C}_{*\mathrm{B}}(Y_*) = \hat{C}_{\mathrm{B}}(Y)A$. This entails that $|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*;\, \gamma_0)\big)| = |A'A|\,|\tilde{\Sigma}\big(\hat{C}(Y;\, \gamma_0)\big)|$ and $|\tilde{\Sigma}_*\big(\hat{C}_{*\mathrm{B}}(Y_*)\big)| = |A'A|\,|\tilde{\Sigma}\big(\hat{C}_{\mathrm{B}}(Y)\big)|$, so that the generalized variance ratios are invariant to the transformation $\tilde{Y}_* = \tilde{Y}A$:

$$\tilde{\Lambda}_*(\gamma_0) = \frac{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*;\, \gamma_0)\big)|}{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)|} = \frac{|\tilde{\Sigma}\big(\hat{C}(Y;\, \gamma_0)\big)|}{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)|} = \tilde{\Lambda}(\gamma_0)\,, \tag{B.4}$$

$$\tilde{\Lambda}_{\mathrm{B}*} = \frac{|\tilde{\Sigma}_*\big(\hat{C}_{*\mathrm{B}}(Y_*)\big)|}{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)|} = \frac{|\tilde{\Sigma}\big(\hat{C}_{\mathrm{B}}(Y)\big)|}{|\tilde{\Sigma}_*\big(\hat{C}(Y)\big)|} = \tilde{\Lambda}_{\mathrm{B}}\,. \tag{B.5}$$

Finally, in view of (2.12) and (6.5), we have

$$LR_*(\gamma_0) = T\ln[\tilde{\Lambda}_*(\gamma_0)] = T\ln[\tilde{\Lambda}(\gamma_0)] = LR(\gamma_0)\,, \tag{B.6}$$

$$LR_{\mathrm{B}*} = T\ln(\tilde{\Lambda}_{\mathrm{B}*}) = T\ln(\tilde{\Lambda}) = LR_{\mathrm{B}}\,. \tag{B.7}$$

□

PROOF OF THEOREM **6.2** Consider a transformation of the form $\tilde{Y}_* = \tilde{Y}K^{-1}$ or, equivalently, $Y_* = YK^{-1} + \tilde{R}_{\mathrm{M}}\iota_n'(I - K^{-1})$. Using (2.1) and (2.2), we then have:

$$Y_* = (XB + WK)K^{-1} + \tilde{R}_{\mathrm{M}}\iota_n'(I - K^{-1}) = XBK^{-1} + \tilde{R}_{\mathrm{M}}\iota_n'(I - K^{-1}) + W$$

$$
\begin{aligned}
&= (\iota_T a' + \tilde{R}_{\mathrm{M}}\beta')K^{-1} + \tilde{R}_{\mathrm{M}}\iota_n'(I - K^{-1}) + W \\
&= \tilde{R}_{\mathrm{M}}\iota_n' + [\iota_T a' + \tilde{R}_{\mathrm{M}}(\beta - \iota_n)']K^{-1} + W \\
&= \tilde{R}_{\mathrm{M}}\iota_n' + X(B - \Delta)K^{-1} + W = \tilde{R}_{\mathrm{M}}\iota_n' + X\bar{B} + W \qquad \text{(B.8)}
\end{aligned}
$$

where $\bar{B} = (B - \Delta)K^{-1}$ and $\Delta = [0,\ \iota_n]'$. In view of the invariance established in Lemma **6.1**, the statistics $LR(\gamma_0)$ and LR_{B} can be viewed as functions of Y_*, which entails that depend on the model parameters (B, K) only through the lower dimensional vector $\bar{B} = (B - \Delta)K^{-1}$. Furthermore, under the null hypothesis, the nuisance parameter only involves γ and $(\beta - \iota_n)'K^{-1}$.

Now the distribution of $LR(\gamma_0)$ and LR_{B} can be explicitly characterized by using (B.4) - (B.5) and observing that

$$
\tilde{\Lambda}(\gamma_0) = \frac{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*;\ \gamma_0)\big)|}{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)|} = \frac{|\hat{W}(\gamma_0)'\hat{W}(\gamma_0)|}{|\hat{W}'\hat{W}|},
$$

$$
\tilde{\Lambda}_{\mathrm{B}} = \frac{|\tilde{\Sigma}_*\big(\hat{C}_{*\mathrm{B}}(Y_*)\big)|}{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)|} = \frac{\inf\{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*;\ \gamma_0)\big)| : \gamma_0 \in \Gamma\}}{|\tilde{\Sigma}_*\big(\hat{C}_*(Y_*)\big)|} = \inf\{\tilde{\Lambda}(\gamma_0) : \gamma_0 \in \Gamma\},
$$

where $\hat{W} = M(X)W$ and

$$
\begin{aligned}
\hat{W}(\gamma_0) &= \bar{M}(\gamma_0)(Y_* - \tilde{R}_{\mathrm{M}}\iota_n') = \bar{M}(\gamma_0)(X\bar{B} + W) \\
&= \bar{M}(\gamma_0)\{[\iota_T a' + \tilde{R}_{\mathrm{M}}(\beta - \iota_n)']K^{-1} + W\} \\
&= \bar{M}(\gamma_0)\{\big[\iota_T\big(a' + \gamma_0(\beta - \iota_n)'\big) + (\tilde{R}_{\mathrm{M}} - \gamma_0\iota_T)(\beta - \iota_n)'\big]K^{-1} + W\} \\
&= \bar{M}(\gamma_0)\{\iota_T\big(a + \gamma_0(\beta - \iota_n)'\big)K^{-1} + W\}.
\end{aligned}
$$

Further, when $\mathcal{H}_{\mathrm{B}}$ holds, we have $a = -\gamma(\beta - \iota_n)$, hence

$$
\hat{W}(\gamma_0) = (\gamma_0 - \gamma)\bar{M}(\gamma_0)\iota_T(\beta - \iota_n)'K^{-1} + \bar{M}(\gamma_0)W.
$$

The theorem then follows on observing that $LR(\gamma_0) = T\ \ln[\tilde{\Lambda}(\gamma_0)]$ and $LR_{\mathrm{B}} = T\ \ln(\tilde{\Lambda}_{\mathrm{B}})$. □

PROOF OF THEOREM **7.1** It is clear that $\mathcal{H}_{\mathrm{B}} = \underset{\gamma_0}{\cup}\mathcal{H}(\gamma_0)$. Since $LR_{\mathrm{B}} = \inf\{LR(\gamma_0) : \gamma_0 \in \Gamma\}$, we have $LR_{\mathrm{B}} \le LR(\gamma_0)$, for any γ_0, hence

$$
\mathsf{P}[LR_{\mathrm{B}} \ge x] \le \mathsf{P}_{(B,\,K)}\big[LR(\gamma_0) \ge x\big],\ \forall x. \qquad \text{(B.9)}
$$

for each γ_0 and for any $(B,\ K)$ compatible with $\mathcal{H}(\gamma_0)$. Furthermore, under $\mathcal{H}_{\mathrm{B}}$, there is a value of γ_0 such that the distribution of $LR(\gamma_0)$ is given by Theorem **4.1**, which entails (7.1). The result for the Gaussian special case then follows upon using (4.2). □

PROOF OF THEOREM **7.2** The result follows from (7.7) and the inequalities $\hat{p}_N^U(\gamma,\ \nu) \le \hat{p}_N^U(\Gamma,\ \nu)$ and $\hat{p}_N^U(\gamma,\ \nu) \le \hat{p}_N^U(\gamma,\ \Omega_{\mathcal{D}}) \le \hat{p}_N^U(\Gamma,\ \Omega_{\mathcal{D}})$. □

PROOF OF THEOREM **7.3** Consider first the case where ν is specified. By (7.6), (2.13) and (4.5), we have:

$$\hat{p}_N^U(\hat{\gamma},\, \nu) \equiv p_N[LR_{\mathrm{B}}^{(0)} \big| \overline{LR}_N(\hat{\gamma},\, \nu)] = p_N[LR^{(0)}(\hat{\gamma}) \big| \overline{LR}_N(\hat{\gamma},\, \nu)] = \hat{p}_N(\hat{\gamma},\, \nu)\,,$$

hence

$$\sup\{\hat{p}_N(\gamma_0,\, \nu) : \gamma_0 \in \Gamma\} \leq \alpha \Rightarrow \hat{p}_N(\hat{\gamma},\, \nu) \leq \alpha \Rightarrow \hat{p}_N^U(\hat{\gamma},\, \nu) \leq \alpha$$

and, on noting that $\sup\{\hat{p}_N(\gamma_0,\, \nu) : \gamma_0 \in \Gamma\} \leq \alpha$ means that $C_\gamma(\alpha,\, \nu)$ is empty,

$$\hat{p}_N^U(\hat{\gamma},\, \nu) > \alpha \Rightarrow \sup\{\hat{p}_N(\gamma_0,\, \nu) : \gamma_0 \in \Gamma\} > \alpha \Rightarrow C_\gamma(\alpha,\, \nu) \neq \emptyset\,,$$

which establishes (7.11). Similarly, for $\nu \in \Omega_{\mathcal{D}}$ unknown,

$$\begin{aligned} \hat{p}_N^U(\hat{\gamma},\, \Omega_{\mathcal{D}}) &= \sup\{\hat{p}_N^U(\hat{\gamma},\, \nu_0) : \nu_0 \in \Omega_{\mathcal{D}}\} = \sup\{p_N[LR_{\mathrm{B}}^{(0)} \big| \overline{LR}_N(\hat{\gamma},\, \nu_0)] : \nu_0 \in \Omega_{\mathcal{D}}\} \\ &= \sup\{p_N[LR^{(0)}(\hat{\gamma}) \big| \overline{LR}_N(\hat{\gamma},\, \nu_0)] : \nu_0 \in \Omega_{\mathcal{D}}\} = \sup\{\hat{p}_N(\hat{\gamma},\, \nu) : \nu_0 \in \Omega_{\mathcal{D}}\}\,, \end{aligned}$$

hence

$$\begin{aligned} \sup\{\hat{p}_N(\gamma_0,\, \nu_0) : \gamma_0 \in \Gamma,\, \nu_0 \in \Omega_{\mathcal{D}}\} \leq \alpha &\Rightarrow \sup\{\hat{p}_N(\hat{\gamma},\, \nu_0) : \nu_0 \in \Omega_{\mathcal{D}}\} \leq \alpha \\ &\Rightarrow \hat{p}_N^U(\hat{\gamma},\, \Omega_{\mathcal{D}}) \leq \alpha\,, \end{aligned}$$

and

$$\hat{p}_N^U(\hat{\gamma},\, \Omega_{\mathcal{D}}) > \alpha \Rightarrow \sup\{\hat{p}_N(\gamma_0,\, \nu_0) : \gamma_0 \in \Gamma,\ \nu_0 \in \Omega_{\mathcal{D}}\} > \alpha \Rightarrow C_\gamma(\alpha;\, \mathcal{D}) \neq \emptyset\,,$$

so that (7.12) is established. □

C. Monte Carlo tests

The Monte Carlo (MC) test procedure goes back to Dwass (1957) and Barnard (1963). Extensions to the nuisance-parameter-dependent case are from Dufour (2006). Here we summarize the underlying methodology (*given a right tailed test with a continuous distribution under the null hypothesis*), as it applies to the test statistics we consider in this paper. For presentation ease, we introduce the following notation. Given a series of $N+1$ variates $S^{(0)},\, S^{(1)},\, \ldots\,,\, S^{(N)}$, let

$$G_N\big[S^{(0)}\,;\, S^{(1)},\, \ldots\,,\, S^{(N)}\big] = \frac{1}{N}\sum_{j=1}^{N} I_{[0,\,\infty)}[S^{(j)} - S^{(0)}]\,,\ I_A[x] = \begin{cases} 1\,, & if\ x \in A\,, \\ 0, & if\ x \notin A\,. \end{cases} \tag{C.1}$$

In what follows, we will use the function $G_N(\,S^{(0)};\, S^{(1)},\, \ldots\,,\, S^{(N)})$ to define empirical type I error probabilities associated with a an observed statistic $S^{(0)}$, given a series of simulated statistics $S^{(1)},\, \ldots\,,\, S^{(N)}$.

First, consider the case:

$$S(y,\, X) = \hat{S}\,(W,\, X)\ ,$$

where W is defined by (2.2), and the distribution of the rows of W is known. Then the conditional distribution of $S(y,\, X)$, given X, is completely determined by the matrix X and the conditional distribution of W given X: the test statistic $S(y,\, X)$ is *pivotal*.

1. Let $S^{(0)}$ be the observed test statistic (based on data).

2. By Monte Carlo methods, draw N *i.i.d.* replications of W :

$$W_{(j)} = [W_1^{(j)},\, \dots\, ,\, W_n^{(j)}]\,, \quad j = 1,\, \dots\, ,\, N\,.$$

3. From each simulated error matrix $W_{(j)}$, compute the statistics

$$\hat{S}(W_{(j)},\, X)\,, \quad j = 1\,,\, \dots\, ,\, N\,.$$

For instance, in the case of the QLR statistic underlying Theorem **4.1**, calculate $T \ln\big(|W_{(j)}'\bar{M}(\gamma_0)W_{(j)}|/|W_{(j)}'MW_{(j)}|\big),\ j = 1\,,\, \dots\, ,\, N\,.$

4. Compute the MC p-value

$$\hat{p}_N\big(S^{(0)}\big) = \frac{NG_N\big[S^{(0)}\,;\, \hat{S}(W_{(1)},\, X)\,,\, \dots\, ,\, \hat{S}(W_{(N)},\, X)\big] + 1}{N+1}\,, \tag{C.2}$$

where the function $G_N(\cdot)$ is as defined in (4.4). Here $NG_N\big[S^{(0)}\,;\, \hat{S}(W_{(1)},\, X),\, \dots\, ,\, \hat{S}(W_{(N)},\, X)\big]$ is the number of simulated values which are greater than or equal to $S^{(0)}$.

5. The MC critical region is

$$\hat{p}_N\big[S^{(0)}\big] \le \alpha\,, \quad 0 < \alpha < 1\,. \tag{C.3}$$

If $\alpha(N+1)$ is an integer, then under the null hypothesis

$$\mathsf{P}\big[\hat{p}_N\big[S^{(0)}\big] \le \alpha\big] = \alpha\,. \tag{C.4}$$

Secondly, suppose that $S(y,\, X)$ is not pivotal but is bounded by a pivotal quantity. Formally, let

$$S(y,\, X) \le \tilde{S}\,(W\,,\, X)$$

where W is defined by (2.2), and the distribution of the rows of W is known. Repeat steps 1-4 above, except that at step 3, obtain draws from the bounding distribution

$$\tilde{S}(W_{(j)}\,,\, X)\,, \quad j = 1\,,\, \dots\, ,\, N\,.$$

To emphasize this distinction (relative to the pivotal statistics case), we refer to the bounds based

p-value in step 3 as $\tilde{p}_N[S]$

$$\tilde{p}_N[S^{(0)}] = \frac{NG_N\big[S^{(0)}\,;\, \tilde{S}(W_{(N)}\,,\, X),\, \ldots\,,\, \tilde{S}(W_{(N)},\, X)\big] + 1}{N+1}\,, \tag{C.5}$$

where the function $G_N(\cdot)$ is as defined in (4.4). The corresponding bound critical region (C.3) is level correct, in these sense that under the null hypothesis

$$\mathsf{P}\big[\hat{p}_N\big[S^{(0)}\big] \leq \alpha\big] \leq \alpha\,.$$

Finally, consider the case where the distribution of W depends on unknown parameters, such as in (2.2) with unknown ν. In this case, given ν, (C.2) and (C.5) yield MC p-values which we will denote respectively $\hat{p}_N[S|\nu]$ and $\tilde{p}_N[S|\nu]$ where the conditioning on ν is emphasized for further reference. Treating ν as a nuisance parameter, the **MMC** (Dufour (2006)) level-correct critical regions corresponds to

$$\sup_{\nu\,\in\,\Phi_0} \hat{p}_N[S|\nu] \leq \alpha \text{ and } \sup_{\nu\,\in\,\Phi_0} \tilde{p}_N[S|\nu] \leq \alpha \tag{C.6}$$

where Φ_0 is a nuisance parameter set consistent with the null hypothesis.

D. Goodness-of-fit tests

The test we propose to invert here is the goodness-of-fit (GF) test introduced in Dufour et al. (2003). Here we summarize the test procedure and refer the reader to Dufour et al. (2003) for proofs, formal algorithms and further references. Consider

$$\mathrm{ESK}(\nu_0) \;=\; \big|\mathrm{SK} - \overline{\mathrm{SK}}(\nu_0)\big|\,, \quad \mathrm{SK} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{i=1}^{T}\hat{d}_{ii}^{3}\,, \tag{D.1}$$

$$\mathrm{EKU}(\nu_0) \;=\; \big|\mathrm{KU} - \overline{\mathrm{KU}}(\nu_0)\big|\,, \quad \mathrm{KU} = \frac{1}{T}\sum_{t=1}^{T}\hat{d}_{tt}^{2}\,, \tag{D.2}$$

where $\hat{d}_{it}$ are the elements of the matrix $\hat{U}(\hat{U}'\hat{U}/T)^{-1}\hat{U}'$ 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). SK and KU are the multivariate skewness and kurtosis criteria introduced by Mardia (1970) in models where the regressor matrix reduces to a vector of ones and considered by Zhou (1993).

$\mathrm{ESK}(\nu_0)$ and $\mathrm{EKU}(\nu_0)$ can be shown to be pivotal under (2.2); thus the MC test technique can be applied to obtain exact p-values [denoted $\hat{p}[\mathrm{ESK}(\nu_0)|\nu_0)$, $\hat{p}[\mathrm{EKU}(\nu_0)|\nu_0)$], *conditional on the same* $\overline{\mathrm{SK}}(\nu_0)$ *and* $\overline{\mathrm{KU}}(\nu_0)$. To obtain a joint test based on these two statistics, we consider the following joint test statistic:

$$\mathrm{CSK} = 1 - \min\left\{\hat{p}[\mathrm{ESK}(\nu_0)|\nu_0),\, \hat{p}[\mathrm{EKU}(\nu_0)|\nu_0)\right\}. \tag{D.3}$$

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

E. Exact diagnostic tests

We consider the tests originally proposed in Dufour, Khalaf and Beaulieu (2002), which consist in applying standard univariate GARCH and variance ratio tests to standardized residuals, namely $\tilde{W}_{it}$, the elements of the *standardized residual* matrix

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

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 $\hat{U}'\hat{U} = S_{\hat{U}}'S_{\hat{U}}$. In particular, we consider: the standard LM-GARCH test statistic [Engle (1982)], its modified version which exploits the one sided nature of the (G)ARCH hypothesis [Lee and King (1993)], and the variance ratio statistic [Lo and MacKinlay (1988, 1989)]. When applied to the standardized residuals of equation i, these statistics can be obtained as follows. Engle's GARCH test statistic, which we will denote E_i is given by TR_i^2, where T is the sample size, R_i^2 is the coefficient of determination in the regression of the equation's squared OLS residuals $\hat{u}_{it}^2$ on a constant and $\hat{u}_{(t-j),i}^2$ $(j = 1, \ldots, q)$. The modified Lee-King statistic is

$$LK_i = \frac{\left\{(T-q)\sum_{t=q+1}^{T}\left[(\tilde{W}_{it}^2/\hat{\sigma}_i^2 - 1)\right]\sum_{j=1}^{q}\tilde{W}_{i,t-j}^2\right\} / \left\{\sum_{t=q+1}^{T}(\tilde{W}_{it}^2/\hat{\sigma}_i^2 - 1)^2\right\}^{1/2}}{\left\{(T-q)\sum_{t=q+1}^{T}\left(\sum_{j=1}^{q}\tilde{W}_{i,t-j}^2\right)^2 - \left(\sum_{t=q+1}^{T}\left(\sum_{j=1}^{q}\tilde{W}_{i,t-j}^2\right)\right)^2\right\}^{1/2}} \tag{E.2}$$

where $\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\tilde{W}_{it}^2$, and the modified variance ratio may be obtained as:

$$VR_i = 1 + 2\sum_{j=1}^{\kappa}\left(1 - \frac{j}{\kappa}\right)\hat{\rho}_{ij}\,, \quad \hat{\rho}_{ij} = \sum_{t=j+1}^{T}\tilde{W}_{it}\tilde{W}_{i,t-j} \,/\, \sum_{t=1}^{T}\tilde{W}_{ti}^2\,. \tag{E.3}$$

In Dufour et al. (2003), we show that under (2.2), $\tilde{W}$ has a distribution which is completely determined by the distribution of W given X, provided K is restricted to be upper triangular. Hence any statistic which depends on the data only through $\tilde{W}$ has a distribution which is invariant to B and Σ, under (2.2). To obtain combined inference across equation, consider the combined statistics:

$$\tilde{E} = 1 - \min_{1\le i\le n}\left[p(\tilde{E}_i)\right], \tag{E.4}$$

$$\widetilde{LK} = 1 - \min_{1\le i\le n}\left[p(\widetilde{LK}_i)\right], \tag{E.5}$$

$$\widetilde{VR} = 1 - \min_{1 \le i \le n} \left[p(\widetilde{VR}_i) \right] , \tag{E.6}$$

where $p(\tilde{E}_i)$, $p(\widetilde{LK}_i)$ and $p(\widetilde{VR}_i)$ refer to p-values; these may be obtained applying a MC test method, or (to cut execution time) using asymptotic null distributions, respectively, $\chi^2(q)$, the standard normal and $N\big[1,\ 2(2\kappa - 1)(\kappa - 1)/(3\kappa)\big]$. Then apply an MMC test procedure (see Appendix C) to the combined statistic imposing (2.2).

References

Allingham, M. (1991), 'Existence theorems in the capital asset pricing model', *Econometrica* **59**, 1169–1174.

Amsler, C. E. and Schmidt, P. (1985), 'A Monte Carlo investigation of the accuracy of multivariate CAPM tests', *Journal of Financial Economics* **14**, 359–375.

Anderson, T. W. (2003), *An Introduction to Multivariate Statistical Analysis*, third edn, John Wiley & Sons, New York.

Banz, R. (1981), 'The relationship between return and market value of common stocks', *Journal of Financial Economics* **5**, 3–18.

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

Beaulieu, M.-C., Dufour, J.-M. and Khalaf, L. (2006), 'Multivariate tests of mean-variance efficiency with possibly non-Gaussian errors: An exact simulation-based approach', *Journal of Business and Economic Statistics* **forthcoming**.

Berk, J. B. (1997), 'Necessary conditions for the CAPM', *Journal of Economic Theory* **73**, 245–257.

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.

Black, F., Jensen, M. C. and Scholes, M. (1972), The capital asset pricing model: Some empirical tests, *in* M. C. Jensen, ed., 'Studies in the Theory of Capital Markets', Vol. 81, Praeger, New York, pp. 79–121.

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

Brennan, M. (1971), 'Capital market equilibrium with divergent borrowing and lending rates', *Journal of Financial Economics* **9**, 3–18.

Buonaccorsi, J. P. and Iyer, H. K. (1984), 'A comparison of confidence regions and designs in estimation of a ratio', *Communications in Statistics, Simulation and Computation* **13**(6), 723–741.

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

Chamberlain, G. (1983), 'A characterization of the distributions that imply mean-variance utility functions', *Journal of Economic Theory* **29**, 185–201.

Chen, N. F., Roll, R. and Ross, S. (1986), 'Economic forces and the stock market', *Journal of Business* **5**, 1–33.

Chou, P.-H. (2000), 'Alternative tests of the zero-beta CAPM', *The Journal of Financial Research* **23**, 469–493.

Dachraoui, K. and Dionne, G. (2003), Conditions ensuring the separability of asset demand by all risk averse investors, Technical report, Risk Management Chair, HEC Montréal, Canada.

Dagenais, M. G. and Dufour, J.-M. (1991), 'Invariance, nonlinear models and asymptotic tests', *Econometrica* **59**, 1601–1615.

Dufour, J.-M. (1989), 'Nonlinear hypotheses, inequality restrictions, and non-nested hypotheses: Exact simultaneous tests in linear regressions', *Econometrica* **57**, 335–355.

Dufour, J.-M. (1990), 'Exact tests and confidence sets in linear regressions with autocorrelated errors', *Econometrica* **58**, 475–494.

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 Jasiak, J. (2001), 'Finite sample limited information inference methods for structural equations and models with generated regressors', *International Economic Review* **42**, 815–843.

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. (2002), Finite-sample diagnostics for multivariate regressions with applications to linear asset pricing models, Technical report, CIRANO and CIREQ, Université de Montréal, and Département d'économique, Université Laval.

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. (2005), Multivariate residual-based finite-sample tests for serial dependence and GARCH with applications to asset pricing models, Technical report, CIRANO and CIREQ, Université de Montréal.

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.

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

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

Fama, E. F. (1965), 'The behaviour of stock prices', *Journal of Business* **60**, 401–424.

Fama, E. F. and French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**, 3–56.

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.

Fama, E. F. and MacBeth, J. (1973), 'Risk, return, and equilibrium: Empirical tests', *Journal of Political Economy* **71**, 607–636.

Ferson, W. E. and Harvey, C. R. (1999), 'Conditioning variables and the cross section of stock returns', *The Journal of Finance* **54**, 1325–1360.

Fieller, E. C. (1940), 'The biological standardization of insulin', *Journal of the Royal Statistical Society (Supplement)* **7**, 1–64.

Fieller, E. C. (1954), 'Some problems in interval estimation', *Journal of the Royal Statistical Society, Series B* **16**, 175–185.

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

Gibbons, M. R., Ross, S. A. and Shanken, J. (1989), 'A test of the efficiency of a given portfolio', *Econometrica* **57**, 1121–1152.

Goffe, W. L., Ferrier, G. D. and Rogers, J. (1994), 'Global optimization of statistical functions with simulated annealing', *Journal of Econometrics* **60**, 65–99.

Harville, D. A. (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag, New York.

Ingersoll, J. (1987), *Theory of Financial Decision Making*, Rowman & Littlefield, Totowa, NJ.

Jagannathan, R. and McGrattan, E. R. (1995), 'The CAPM debate', *Federal Reserve Bank of Minneapolis Quarterly Review* **19**(4), 2–17.

Jagannathan, R. and Wang, Z. (1996), 'The conditional CAPM and the cross section of expected returns', *The Journal of Finance* **51**, 3–53.

Jobson, J. and Korkie, B. (1982), 'Potential performance and tests of portfolio efficiency', *Journal of Financial Economics* **10**, 433–466.

Kandel, S. (1984), 'The likelihood ratio test statistics of mean-variance efficiency without a riskless asset', *Journal of Financial Economics* **13**, 575–592.

Kandel, S. (1986), 'The geometry of the likelihood estimator of the zero beta return', *Journal of Finance* **31**, 339–346.

Kandel, S. and Stambaugh, R. F. (1989), 'A mean-variance framework for tests of asset pricing models', *The Review of Financial Studies* **2**, 125–156.

Kariya, T. (1985), *Testing in the Multivariate General Linear Model*, number 22 *in* 'Economic Research Series, The Institute of Economic Research, Hitotsubashi University, Japan', Kinokuniya Company Ltd., Tokyo.

Kim, D. (1995), 'The errors in the variables problem in the cross-section of expected stock returns', *Journal of Finance* **14**, 1605–1634.

Kleibergen, F. (2002), 'Pivotal statistics for testing structural parameters in instrumental variables regression', *Econometrica* **70**(5), 1781–1803.

Koschat, M. A. (1987), 'A characterization of the Fieller solution', *The Annals of Statistics* **15**, 462–468.

Krishnaiah, P. R., ed. (1980), *Handbook of Statistics 1: Analysis of Variance*, North-Holland, Amsterdam.

Lee, J. H. and King, M. L. (1993), 'A locally most mean powerful based score test for ARCH and GARCH regression disturbances', *Journal of Business and Economic Statistics* **11**, 17–27. Correction 12 (1994), 139.

Litzenberger, R. and Ramaswamy, K. (1979), 'The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence', *Journal of Financial Economics* **7**, 163–195.

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.

Lo, A. and MacKinlay, C. (1989), 'The size and power of the variance ratio test in finite samples: A Monte Carlo investigation', *Journal of Econometrics* **40**, 203–238.

MacBeth, J. (1979), Tests of the Two Parameter Model of Capital Market Equilibrium, PhD thesis, University of Chicago, Chicago, Illinois.

MacKinlay, A. C. (1987), 'On multivariate tests of the Capital Asset Pricing Model', *Journal of Financial Economics* **18**, 341–372.

MacKinlay, A. C. and Richardson, M. P. (1991), 'Using generalized method of moments to test mean-variance efficiency', *The Journal of Finance* **46**, 511–527.

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

Moreira, M. J. (2003), 'A conditional likelihood ratio test for structural models', *Econometrica* **71**(4), 1027–1048.

Nielsen, L. T. (1990), 'Existence of equilibrium in CAPM', *Journal of Economic Theory* **52**, 223–231.

Pagan, A. (1984), 'Econometric issues in the analysis of regressions with generated regressors', *International Economic Review* **25**, 221–247.

Perold, A. F. (2004), 'The Capital Asset Pricing Model', *The Journal of Economic Perspectives* **18**(3), 3–24.

Rao, C. R. (1973), *Linear Statistical Inference and its Applications*, second edn, John Wiley & Sons, New York.

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

Roll, R. (1985), 'A note on the geometry of Shanken's CSR T^2 test for mean/variance efficiency', *Journal of Financial Economics* **14**, 349–357.

Ross, S. A. (1978), 'Mutual fund separation in financial theory - The separating distributions', *Journal of Economic Theory* **17**, 254–286.

Samuelson, P. (1967), 'On multivariate tests of the Capital Asset Pricing Model', *Journal of Financial and Quantitative Analysis* **2**, 107–122.

Schott, J. R. (1984), 'Optimal bounds for the distribution of some test criteria for tests of dimensionality', *Biometrika* **71**, 561–567.

Seber, G. A. F. (1984), *Multivariate Observations*, John Wiley & Sons, New York.

Shanken, J. (1985), 'Multivariate tests of the zero-beta CAPM', *Journal of Financial Economics* **14**, 325–348.

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

Shanken, J. (1990), 'Intertemporal asset pricing: An empirical investigation', *Journal of Econometrics* **45**, 99–120.

Shanken, J. (1992), ‘On the estimation of beta-pricing models’, *Review of Financial Studies* **5**, 1–33.

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.

Staiger, D. and Stock, J. H. (1997), ‘Instrumental variables regression with weak instruments’, *Econometrica* **65**(3), 557–586.

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

Stock, J. H. and Wright, J. H. (2000), ‘GMM with weak identification’, *Econometrica* **68**, 1097–1126.

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.

Wang, J. and Zivot, E. (1998), ‘Inference on structural parameters in instrumental variables regression with weak instruments’, *Econometrica* **66**(6), 1389–1404.

Zerbe, G. O. (1978), ‘On Fieller’s theorem and the general linear model’, *The American Statistician* **32**(3), 103–105.

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

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

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

Zivot, E., Startz, R. and Nelson, C. R. (1998), ‘Valid confidence intervals and inference in the presence of weak instruments’, *International Economic Review* **39**, 1119–1144.