Testing Asset Pricing Models with Coskewness.

Giovanni Barone Adesi
(Università della Svizzera Italiana, Lugano, CH,
and
Cass Business School, London, U.K.)

Patrick Gagliardini
(Università della Svizzera Italiana, Lugano, CH)

Giovanni Urga * †
(Ca Business School, London, U.K.)

January 30, 2003

* *Corresponding author:* Giovanni Urga, Faculty of Finance, Cass Business School, Frobisher Crescent, Barbican Centre, London EC2Y 8HB (U.K.). Tel.: +/44/20/70408696, Fax.: ++/44/20/70408881, e-mail: g.urga@city.ac.uk http://www.staff.city.ac.uk/~giourga/

† We wish to thank participants to the 9th Conference on Panel Data, the 10th annual conference of European Financial Management Association, and the conference on "Multimoment Capital Asset Pricing Models and Related Topics" for helpful discussions on an earlier version of this paper. C. R. Harvey, R. Jagannathan, R. Morck, M. Rockinger, T. Wansbeek, and C. Zhang provided very useful comments which helped to improve the paper. We wish to thank K. French, R. Jagannathan and R. Kan for providing us with their data set. Swiss NCCR Finrisk is gratefully acknowledged by the first two authors. The usual disclaimer applies.

Testing Asset Pricing Models with Coskewness

Abstract

In this paper we investigate portfolio coskewness using a quadratic market model as return generating process. It is shown that portfolios of small (large) firms have negative (positive) coskewness with market. An asset pricing model including coskewness is tested through the restrictions it imposes on the return generating process. We find evidence of an additional component in portfolios expected excess returns, which is not explained by neither covariance nor coskewness with the market. However, this unexplained component is constant across portfolios in our sample, and modest in magnitude. We investigate the implications of erroneously neglecting coskewness for testing asset pricing models, with particular interest for the empirically detected explanatory power of size.

Key Words: Coskewness, Asset Pricing Models, Factor Models, Statistical Tests

JEL Classification: C12, C32, C52, G12

Asset pricing models generally express expected returns on financial assets as linear functions of covariances of returns with some systematic risk factors. Sharpe (1964), Lintner (1965), Black (1972), Merton (1973), Kraus and Litzenberger (1976), Ross (1976), Breeden (1979), Barone Adesi and Talwar (1983), Barone Adesi (1985), Jagannathan and Wang (1996), Harvey and Siddique (1999,2000), Dittmar (2002) have proposed several formulations of this general paradigm. However, most of the empirical tests proposed to date have produced negative or ambiguous results. These findings have spurred renewed interest in the statistical properties of testing methodologies currently available. Among recent studies, Shanken (1992) and Kan and Zhang (1999a,b) provide analyses of the statistical methodologies commonly employed and highlight the sources of ambiguity that plague their findings.

Although a full specification of the return generating process is not needed for the formulation of most asset pricing models, it appears that only its preliminary knowledge may lead to the design of reliable tests. Because this condition is never met in practice, researchers are forced to make unpalatable choices between two alternative approaches. On the one hand, powerful tests can be designed in the context of a (fully) specified return generating process, but they are misleading in the presence of possible model misspecifications. On the other hand, more tolerant tests may be considered, but they may lack of power, as noted by Kan and Zhou (1999a,b) and Jagannathan and Wang (2001). Notice that the first choice may lead not only to the rejection of correct models, but also to the acceptance of irrelevant factors as sources of systematic risk, as noted by Kan and Zhang (1999a,b).

To complicate the picture, a number of empirical regularities have been detected. Among them, Banz (1981) relates expected returns to firm size, Fama and French (1995) link expected returns also to the ratio of book to market value. Although the persistence of these anomalies over time is still subject to debate, the evidence suggests that the mean-variance CAPM is not a satisfactory descritpion of market equilibrium.

Pricing anomalies may be related to the possibility that useless factors appear to be priced. Of course it is also possible that pricing anomalies proxy for omitted factors. While statistical tests do not allow us to choose among these two possible explanations of pricing anomalies, Kan and Zhang (1999a,b) suggest that perhaps large increase in R^2 and persistence of sign and size of coefficients over time are most likely to be associated with truly priced factors.

In the light of the considerations above, the main aim of this paper is to consider coskewness and its role in testing asset pricing models, using a data set of monthly returns on 10 stock portfolios. Following Harvey and Siddique (2000), an asset is defined to have "positive coskewness" with the market when the residuals of the regression of its returns on a constant and the market returns are positively correlated with squared market returns. Therefore, an asset with positive (negative) coskewness decreases (increases) the risk of the portfolio to large absolute market returns, and should command a lower (higher) expected return in equilibrium.

Kraus and Litzenberger (1976), Barone-Adesi (1985) and Harvey and Siddique (2000) have studied non-normal[1] asset pricing models related to coskewness. Kraus and Litzenberger (1976) and Harvey and Siddique (2000) formulate expected returns as function of covariance and coskewness with the market portfolio. In particular, Harvey and Siddique (2000) assess the importance of coskewness for explaining assets expected returns by the increase of R^2 in cross-sectional regressions. More recently, Dittmar (2002) presents a framework in which agents are also adverse to kurtosis, implying that asset returns are influenced by both coskewness and cokurtosis with the return on aggregate wealth. He tests this extended asset pricing model within a Generalized Method of Moment (GMM) framework [see Hansen (1982)]. Their formulations are very general, since the specification of an underlying return generating process is not required. However, we are concerned about the possible lack of power of these methodologies, which is worsened in this context by the fact that covariance and coskewness with market are almost perfectly collinear across portfolios. To remedy that, in this paper we propose a formulation [see also Barone-Adesi (1985)] which assumes the quadratic market model as the return generating process. The quadratic market model is an extension of the traditional market model [Sharpe (1964), Lintner (1965)], including the square of the market returns as an additional factor. The coefficients of the quadratic factor are the coskewness coefficients of the portfolios. Since market returns and the square of the market returns are almost orthogonal regressors, we obtain a precise test for the significancy of coskewness coefficients. In addition, this framework allows us to test an asset pricing model with coskewness by testing the restrictions which it imposes on the coefficients of the quadratic market model [2]. The specification of a return generating process provides tests of superior power as confirmed in a series of Monte Carlo simulations (see Section IV).

In addition to evaluate asset pricing models which include coskewness, it is also important to investigate the consequences on asset pricing tests when coskewness is erroneously neglected. We consider the possibility that portfolio characteristics such as size are empirically found to explain expected excess returns since a truly priced factor (coskewness) is omitted. To explain this, let us assume that coskewness is truly priced, but it is omitted in as asset pricing model. Then, if market coskewness is correlated with a variable such as size, this variable will have spurious explanatory power for the cross-section of expected returns, since it proxies for omitted coskewness. In our empirical application (see Section III) we actually find that coskewness and size are correlated. This suggests that a possible explanation for the empirically observed relation between size and assets excess returns is the omission of a systematic risk factor, namely *market coskewness* [3] [4].

The remaining of the paper is organized as follows. Section I introduces the quadratic market model. An asset pricing model including coskewness is derived from it using arbitrage pricing, and various testing methodologies are discussed. Section II reports estimators and test statistics used in the empirical part of the paper. Section III describes the data, and reports empirical results. Section IV provides Monte Carlo simulations for investigating the finite sample properties of the test statistics, and Section V concludes.

I Asset Pricing Models with Coskewness.

In this section we introduce the econometric formulations which are considered in this paper. In turn, we describe the return generating process (I.A), we derive the corresponding restricted equilibrium models (I.B), and finally compare our approach with a GMM framework (I.C).

I.A The Quadratic Market Model

Factor models are amongst the most widely used return generating processes in financial econometrics. They explain comovements in asset returns as arising from the common effect of a (small) number of underlying variables, called factors [see e.g. Campbell, Lo, MacKinlay (1987) and Gourieroux, Jasiak (2001)]. In this paper, a linear two-factor model, called quadratic market model, is used as return generating process. Market returns and the square of the market returns are its two factors. Specifically, let us denote by R_t the $N \times 1$ vector of returns in period t of N portfolios, and by $R_{M,t}$ the return of the market. If $R_{F,t}$ is the return in period t of a (conditionally) risk free asset, excess returns are defined as: $r_t = R_t - R_{F,t}\iota$, $r_{M,t} = R_{M,t} - R_{F,t}$, $q_{M,t} = R^2_{M,t} - R_{F,t}$, where ι is a $N \times 1$ vector of ones. The quadratic market model is then specified by:

$$\begin{array}{ll} & r_t = \alpha + \beta r_{M,t} + \gamma q_{M,t} + \varepsilon_t,\ t = 1, ..., T, \\ \mathcal{H}_F : & \gamma \neq 0 \end{array} \tag{1}$$

where α is a $N \times 1$ vector of intercepts, β and γ are $N \times 1$ vectors of sensitivities and ε_t is an $N \times 1$ vector of errors satisfying[5]:

$$E\left[\varepsilon_t \mid \underline{R_{M,t}}, \underline{R_{F,t}}\right] = 0.$$

The quadratic market model is a direct extension of the well-known market model [Sharpe (1964), Lintner (1965)], which corresponds to the restriction $\gamma = 0$ in (1):

$$\begin{array}{ll} & r_t = \alpha + \beta r_{M,t} + \varepsilon_t,\ t = 1, ..., T, \\ \mathcal{H}^*_F : & \gamma = 0 \text{ in } (1). \end{array} \tag{2}$$

The motivation for including the square of the market returns is to fully account for coskewness with the market portfolio. In fact, deviations from the linear relation between asset returns and market returns implied by (2) are empirically observed. Indeed, for some classes of assets, residuals from the regression of returns on a constant and market returns tend to be positively (negatively) correlated with squared market returns. These assets show therefore a tendency to have relatively higher (lower) returns when the market experiences high absolute returns, and are said to have positive (negative) coskewness with the market. This is exactly what we find in our empirical investigations (see Section III), when in accordance with the results of Harvey and Siddique (2000) we find that portfolios formed by assets of small firms tend to have a negative coskewness with the market, whereas portfolios formed by assets of large firms have positive market coskewness. In addition to classical beta, market coskewness is therefore another very important risk characteristic: an asset that has positive coskewness with the market diminishes the risk of the portfolio with respect to large absolute market returns, and, everything else being equal, investors should prefer assets with positive market coskewness to those with negative coskewness. The quadratic market model (1) is a specification which provides us with a very simple way to take into account market coskewness. Indeed, we have:

$$\gamma = \frac{1}{V\left[\epsilon_{q,t}\right]} cov\left[\epsilon_t, R_{M,t}^2\right], \tag{3}$$

where ϵ_t ($\epsilon_{q,t}$) are the residuals from a theoretical regression of portfolio returns R_t (market square returns $R_{M,t}^2$) on a constant and market returns $R_{M,t}$ [6]. We use the estimate of γ in model (1) to investigate the properties of the coskewness coefficients of the N portfolios. The statistical (joint) significancy of coskewness γ is assessed by testing the null hypothesis $\mathcal{H}_F^*$ against the alternative $\mathcal{H}_F$.

I.B Restricted equilibrium models

From the point of view of financial economics, a linear factor model is only a return generating process, which is not necessarily consistent with notions of economic equilibrium. Constraints on its coefficients are imposed e.g. by arbitrage pricing [Ross (1976), and Chamberlain and Rothschild (1983)]. The arbitrage pricing theory (APT) implies that expected excess returns of assets following the factor model (1) satisfy the restriction[7] [Barone-Adesi (1985)]:

$$E(r_t) = \beta\lambda_1 + \gamma\lambda_2, \tag{4}$$

where λ_1 and λ_2 are expected excess returns on portfolios whose excess returns are perfectly correlated with $r_{M,t}$ and $q_{M,t}$ respectively. Equation (4) is in the form of a typical linear asset pricing model, which relates expected excess returns to covariances and coskewnesses to market. In this paper we test the asset pricing model with coskewness (4) through the restrictions it imposes on the coefficients of the return generating process (1). Let us derive these restrictions. Since the excess market return $r_{M,t}$ satisfies (4), it must be that

$$\lambda_1 = E(r_{M,t}). \tag{5}$$

A similar restriction doesn't hold for the second factor since it is not a traded asset. However, we expect $\lambda_2 < 0$, since assets with positive coskewness decrease the risk of a portfolio with respect to large absolute market returns, and therefore should command a lower risk premium in an arbitrage equilibrium. By taking expectations on both sides of (1), and substituting (4) and (5), we deduce that the asset pricing model (4) implies the cross-equation restriction $\alpha = \vartheta\gamma$, where ϑ is the scalar parameter $\vartheta = \lambda_2 - E(q_{M,t})$. Thus arbitrage pricing is consistent with the following restricted model:

$$\begin{array}{rl} & r_t = \beta r_{M,t} + \gamma q_{M,t} + \gamma\vartheta + \varepsilon_t,\ t = 1, ..., T, \\ \mathcal{H}_1: & \exists\vartheta : \alpha = \vartheta\gamma \text{ in (1).} \end{array} \quad (6)$$

Therefore, the asset pricing model with coskewness (4) is tested by testing $\mathcal{H}_1$ against $\mathcal{H}_F$.

If (4) turns out not to be supported by the data, this implies the existence of an additional component $\widetilde{\alpha}$, a $N \times 1$ vector, in expected excess returns, other than those related to market risk and coskewness risk: $E(r_t) = \beta\lambda_1 + \gamma\lambda_2 + \widetilde{\alpha}$. In this case, the intercepts α of model (1) satisfies the restriction: $\alpha = \vartheta\gamma + \widetilde{\alpha}$. It is crucial to investigate how the additional component $\widetilde{\alpha}$ varies across assets. Indeed, if this component arises from an omitted factor, it will provide us with information about the sensitivities of portfolios to this factor. Furthermore, variables representing portfolio characteristics, which turn out to be correlated with $\widetilde{\alpha}$ across portfolios, will have spurious explanatory power for expected excess returns, since they proxy for the sensitivities of the omitted factor. A case of particular interest arises when $\widetilde{\alpha}$ is homogeneous across assets: $\widetilde{\alpha} = \lambda_0\iota$, where λ_0 is a scalar, that is:

$$E(r_t) = \iota\lambda_0 + \beta\lambda_1 + \gamma\lambda_2, \quad (7)$$

corresponding to the following specification:

$$\begin{array}{rl} & r_t = \beta r_{M,t} + \gamma q_{M,t} + \gamma\vartheta + \lambda_0\iota + \varepsilon_t,\ t = 1, ..., T, \\ \mathcal{H}_2: & \exists\vartheta, \lambda_0 : \alpha = \vartheta\gamma + \lambda_0\iota \text{ in (1).} \end{array} \quad (8)$$

Specification (8) corresponds to the case where the factor omitted in model (4) has homogeneous sensitivities across portfolios. From (7), λ_0 may be interpreted as the expected excess returns of a portfolios with zero covariance and coskewness with the market. Such a portfolio may correspond to the analogous of the zero-beta portfolio in the Black version of the Capital Asset Pricing Model (CAPM) [Black (1972)]. Alternatively, $\lambda_0 > 0$ ($\lambda_0 < 0$) may be due to the use of a risk-free rate lower (higher) than the actual rate investors face. With reference to the observed empirical regularities and model misspecifications, the importance of model (8) is that, if hypothesis $\mathcal{H}_2$ is not rejected against $\mathcal{H}_F$, we expect portfolio characteristics such as size not to have additional explanatory power for expected excess returns, once that coskewness is taken into account. In addition, a more powerful evaluation of the validity of the asset pricing model (4) should be provided by a test of $\mathcal{H}_1$ against the alternative $\mathcal{H}_2$.

I.C The GMM framework

Asset pricing models of the type (4) are considered in Kraus and Litzenberger (1976) and Harvey and Siddique (2000). Harvey and Siddique (2000) introduce their specification as a model where the stochastic discount factor is quadratic in market returns. Specifically, in our notation, equation (4) is equivalent to the orthogonality condition:

$$E\left[r_t m_t\left(\delta\right)\right]=0, \tag{9}$$

where the stochastic discount factor $m_t(\delta)$ is given by: $m_t(\delta) = 1 - r_{m,t}\delta_1 - q_{m,t}\delta_2$, and $\delta = (\delta_1, \delta_2)$ is a two-dimensional parameter. A quadratic stochastic discount factor $m_t(\delta)$ can be justified as a (formal) second order Taylor expansion of a stochastic discount factor, which is nonlinear in the market returns. Thus, in this approach, the derivation and the testing of (9) do not require a prior specification of a data generating process. More recently, in a similar conditional GMM framework, Dittmar (2002) uses a stochastic discount factor model embodying both quadratic and cubic terms, and the validity of the model is tested by a GMM statistics using the weighting matrix proposed in Jagannathan and Wang (1996) and Hansen and Jagannathan (1997). As explained earlier in the paper, the main feature of our paper, with respect Harvey and Siddique (2000) and Dittmar (2002) contributions, is that we focus on testing the asset pricing model with coskewness (4) through the restrictions it imposes on the return generating process (1), instead of adopting a methodology using an unspecified alternative (e.g. by a GMM test).

II Estimators and Test Statistics.

This section derives the estimators and test statistics used in our empirical applications. We consider various procedures widely investigated in the literature [see e.g. Campbell, Lo, MacKinlay (1997) and Gourieroux, Jasiak (2001)], and derive their properties within the alternative coskewness asset pricing models. For completeness, and only when necessary, full derivations are provided in the Appendices.

We assume that the error term ε_t in (1) with $t = 1, \ldots.., T$, is an homoscedastic martingale difference sequence satisfying:

$$\begin{aligned} E\left[\varepsilon_t|\underline{\varepsilon_{t-1}}, \underline{R_{M,t}}, \underline{R_{F,t}}\right] &= 0, \\ E\left[\varepsilon_t\varepsilon_t^{'}|\underline{\varepsilon_{t-1}}, \underline{R_{M,t}}, \underline{R_{F,t}}\right] &= \Sigma, \end{aligned} \tag{10}$$

where Σ is a positive definite $N \times N$ matrix. The factor $f_t = (r_{M,t}, q_{M,t})^{'}$ is supposed to be exogenous in the sense of Engle, Hendry and Richard (1988), and we denote by μ and Σ_f its expectation and variance-covariance matrix, respectively. We conduct estimation and inference in the framework of Pseudo Maximum Likelihood (PML) methods [White (1981), Gourieroux, Monfort and Trognon (1984), Bollerslev and Wooldridge (1992)]. If θ denotes the parameter of interest in the model under consideration, the PML estimator is defined by the maximization:

$$\widehat{\theta} = \arg\max_{\theta} L_T(\theta), \tag{11}$$

where the criterium $L_T(\theta)$ is a (conditional) pseudo-loglikelihood, i.e. the (conditional) loglikelihood of the model, assuming a given conditional distribution for ε_t satisfying (10) and such that the resulting pseudo true density of the model is exponential quadratic. Under regularity assumptions, the PML estimator $\widehat{\theta}$ is consistent, for any chosen conditional distribution of ε_t satisfying the above conditions (see above references). $\widehat{\theta}$ is efficient when the pseudo conditional distribution of ε_t coincides with the true one, being then the PML estimator identical with the maximum likelihood (ML) estimator. Since the PML estimator is based on the maximization of a statistical criterion, hypothesis testing can be conducted by usual general asymptotic tests. In what follows, we will systematically analyze, along those lines, the alternative specifications introduced in Section I.

II.A The return generating process

The quadratic market model (1) [and the market model (2)] are Seemingly Unrelated Regressions (SUR) systems [Zellner (1962)], with the same regressors in each equation. Denoting by θ the parameters[8] of interest in model (1):

$$\theta = (\alpha^{'}, \beta^{'}, \gamma^{'}, vech(\Sigma)^{'})^{'},$$

the PML estimator of θ based on the normal family is obtained by maximizing:

$$L_T(\theta) = -\frac{T}{2}\log\det\Sigma - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t(\theta)'\Sigma^{-1}\varepsilon_t(\theta), \tag{12}$$

where

$$\varepsilon_t(\theta) = r_t - \alpha - \beta r_{M,t} - \gamma q_{M,t}, \quad t = 1, ..., T.$$

As is well-known, the PML estimator for $(\alpha', \beta', \gamma')'$ is equivalent to the GLS estimator on the SUR system and also to the OLS estimator performed equation by equation in (1). Let B denote the $N \times 3$ matrix defined by $B = [\alpha\ \beta\ \gamma]$. The asymptotic distribution of the PML estimator $\widehat{B} = \left[\widehat{\alpha}\ \widehat{\beta}\ \widehat{\gamma}\right]$ is given by:

$$\sqrt{T}\left(\widehat{B} - B\right) \overset{d}{\longrightarrow} N(0, \Sigma \quad E\left[F_t F_t'\right]^{-1}), \tag{13}$$

where $F_t = (1, r_{M,t}, q_{M,t})'$.

Let us now consider the (joint) significancy of the coskewness coefficients by testing $\mathcal{H}_F^*$ against $\mathcal{H}_F$: $\gamma = 0$. It can be easily performed by a Wald statistics, which is given by[9] [see (13)]:

$$\xi_T^{F*} = T\frac{1}{\widehat{\Sigma}_f^{22}}\widehat{\gamma}'\widehat{\Sigma}^{-1}\widehat{\gamma}. \tag{14}$$

Statistics ξ_T^{F*} is asymptotically $\chi^2(p)$-distributed, with $p = N$, when $T \longrightarrow \infty$.

II.B Restricted equilibrium models

Let us now consider the constrained models (6) and (8) derived by arbitrage equilibrium. These models are more complicated since they entail cross-equation restrictions. We denote by:

$$\theta = \left(\beta', \gamma', \vartheta, \lambda_0, vech\,(\Sigma)'\right)',$$

the vector of parameters of model (8). The PML estimator of θ based on a normal pseudo conditional loglikelihood is defined by maximization of:

$$L_T\,(\theta) = -\frac{T}{2}\log\det\Sigma - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t(\theta)'\Sigma^{-1}\varepsilon_t(\theta), \tag{15}$$

where:

$$\varepsilon_t(\theta) = r_t - \beta r_{M,t} - \gamma q_{M,t} - \gamma\vartheta - \lambda_0 \iota, \quad t = 1, ..., T.$$

The PML estimator is given by the following system of implicit equations [see Appendix A]:

$$\left(\widehat{\beta}', \widehat{\gamma}'\right)' = \left(\sum_{t=1}^{T} (r_t - \widehat{\lambda}_0 \iota)\widehat{H}_t'\right)\left(\sum_{t=1}^{T} \widehat{H}_t \widehat{H}_t'\right)^{-1}, \tag{16}$$

$$\left(\widehat{\vartheta}, \widehat{\lambda}_0\right)' = (\widehat{Z}'\widehat{\Sigma}^{-1}\widehat{Z})^{-1}\widehat{Z}'\widehat{\Sigma}^{-1}\left(\overline{r} - \widehat{\beta}\overline{r}_M - \widehat{\gamma}\overline{q}_M\right), \tag{17}$$

$$\widehat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}_t', \tag{18}$$

where:

$$\widehat{\varepsilon}_t = r_t - \widehat{\beta} r_{M,t} - \widehat{\gamma} q_{M,t} - \widehat{\gamma}\widehat{\vartheta} - \widehat{\lambda}_0 \iota,$$

$$\widehat{H}_t = \left(r_{M,t}, q_{M,t} + \widehat{\vartheta}\right)', \quad \widehat{Z} = (\widehat{\gamma}, \iota),$$

and $\overline{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$, $\overline{r}_M = \frac{1}{T}\sum_{t=1}^{T} r_{M,t}$, $\overline{q}_M = \frac{1}{T}\sum_{t=1}^{T} q_{M,t}$. An estimator for $\lambda = (\lambda_1, \lambda_2)'$ is simply obtained by:

$$\widehat{\lambda} = \widehat{\mu} + \begin{pmatrix} 0 \\ \widehat{\vartheta} \end{pmatrix}. \tag{19}$$

Note that $\left(\widehat{\beta}', \widehat{\gamma}'\right)'$ is obtained by (time series) OLS regressions of $r_t - \widehat{\lambda}_0 \iota$ on $\widehat{H}_t$ in a SUR system, performed equation by equation, whereas $\left(\widehat{\vartheta}, \widehat{\lambda}_0\right)'$ is obtained by (cross-sectional) GLS regression of $\overline{r} - \widehat{\beta}\overline{r}_M - \widehat{\gamma}\overline{q}_M$ on $\widehat{Z}$. A step of a feasible algorithm consists in: a) starting from old estimates; b) computing $\left(\widehat{\beta}', \widehat{\gamma}'\right)'$ from (16); c) computing $\left(\widehat{\vartheta}, \widehat{\lambda}_0\right)'$ from (17) using new estimates for $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{Z}$; d) computing $\widehat{\Sigma}$ from (18), using new estimates. The procedure is iterated until a convergence criterion is met. The starting values for β, γ and Σ are provided by the unrestricted estimates on model (1), whereas for the parameters λ_0 and ϑ they are provided by equation (17) [where estimates from (1) are used]. The asymptotic distributions of the PML estimator are reported in Appendix A. In particular, it is shown that the asymptotic variance of the estimator of $\left(\beta^{'}, \gamma^{'}, \vartheta, \lambda_0, \lambda_1, \lambda_2\right)$ is independent of the true distribution of the error term ε_t, as long as this satisfies the conditions for PML estimation. The results for constrained PML estimation of models (6) follow by setting $\lambda_0 = 0$, $\widehat{Z} = \widehat{\gamma}$, and deleting the vector ι.

Let us now consider testing hypotheses $\mathcal{H}_1$ and $\mathcal{H}_2$ against the alternative $\mathcal{H}_F$. If θ denotes the parameter of model (1), these hypotheses can be written in mixed form:

$$\{\theta : \exists a \in A \subset \mathbb{R}^q : g(\theta, a) = 0\}, \tag{20}$$

where g is a vector function with values in $\mathbb{R}^r$. Assuming that the rank conditions:

$$rank\left(\frac{\partial g}{\partial \theta^{'}}\right) = r, \quad rank\left(\frac{\partial g}{\partial a^{'}}\right) = q,$$

are satisfied at the true values θ^0, a^0, a specification test for the hypothesis (20) based on Asymptotic Least Squares (ALS) consists in verifying whether the constraints $g(\widehat{\theta}, a) = 0$ are satisfied, where $\widehat{\theta}$ is an unconstrained estimator of θ (the PML estimator in our case) [Gourieroux, Monfort and Trognon (1985)]. It is based on the following statistics:

$$\xi_T = \arg\min_a T g(\widehat{\theta}, a)^{'} \widehat{S} g(\widehat{\theta}, a),$$

where $\widehat{S}$ is a consistent estimator for

$$S_0 = \left(\frac{\partial g}{\partial \theta^{'}} \;_0 \frac{\partial g^{'}}{\partial \theta}\right)^{-1},$$

evaluated at the true values θ^0, a^0, where $\;_0 = V_{as}\left[\sqrt{T}\left(\widehat{\theta} - \theta^0\right)\right]$. Under regularity conditions, ξ_T is asymptotically $\chi^2(r-q)$-distributed, and is asymptotically equivalent to the other asymptotic tests[10].

We report the ALS test statistics for testing the hypotheses $\mathcal{H}_2$ and $\mathcal{H}_1$ against the alternative $\mathcal{H}_F$ [they are fully derived in Appendix B]. The hypothesis $\mathcal{H}_1$ against $\mathcal{H}_F$ is tested by the statistics:

$$\xi_T^1 = T \frac{\left(\widehat{\alpha} - \widetilde{\vartheta}\widehat{\gamma}\right)^{'} \widehat{\Sigma}^{-1} \left(\widehat{\alpha} - \widetilde{\vartheta}\widehat{\gamma}\right)}{1 + \widetilde{\lambda}^{'} \widehat{\Sigma}_f^{-1} \widetilde{\lambda}} \sim \chi^2(p), \tag{21}$$

with $p = N - 1$, where $\widetilde{\lambda} = \widehat{\mu} + \left(0, \widetilde{\vartheta}\right)'$, and:

$$\begin{aligned} \widetilde{\vartheta} &= \arg\min_{\vartheta} \left(\widehat{\alpha} - \vartheta\widehat{\gamma}\right)' \widehat{\Sigma}^{-1} \left(\widehat{\alpha} - \vartheta\widehat{\gamma}\right) \\ &= (\widehat{\gamma}'\widehat{\Sigma}^{-1}\widehat{\gamma})^{-1}\widehat{\gamma}'\widehat{\Sigma}^{-1}\widehat{\alpha}. \end{aligned}$$

The hypothesis $\mathcal{H}_2$ against $\mathcal{H}_F$ is tested by the statistics:

$$\xi_T^2 = T \frac{\left(\widehat{\alpha} - \widetilde{\vartheta}\widehat{\gamma} - \widetilde{\lambda}_0\iota\right)' \widehat{\Sigma}^{-1} \left(\widehat{\alpha} - \widetilde{\vartheta}\widehat{\gamma} - \widetilde{\lambda}_0\iota\right)}{1 + \widetilde{\lambda}' \widehat{\Sigma}_f^{-1} \widetilde{\lambda}} \sim \chi^2(p), \tag{22}$$

with $p = N - 2$, where $\widetilde{\lambda} = \widehat{\mu} + \left(0, \widetilde{\vartheta}\right)'$, and:

$$\begin{aligned} \left(\widetilde{\vartheta}, \widetilde{\lambda}_0\right)' &= \arg\min_{\vartheta, \lambda_0} \left(\widehat{\alpha} - \vartheta\widehat{\gamma} - \lambda_0\iota\right)' \widehat{\Sigma}^{-1} \left(\widehat{\alpha} - \vartheta\widehat{\gamma} - \lambda_0\iota\right) \\ &= (\widehat{Z}'\widehat{\Sigma}^{-1}\widehat{Z})^{-1}\widehat{Z}'\widehat{\Sigma}^{-1}\widehat{\alpha}, \qquad \widehat{Z} = \left(\widehat{\gamma}, \iota\right). \end{aligned}$$

Finally a test of $\mathcal{H}_1$ against $\mathcal{H}_2$ is simply performed by a t-test for the parameter λ_0.

III Empirical Results.

In this section we report the results of our tests: we estimate the quadratic market model (1), and test asset pricing models with coskewness (6) and (8). We begin with a brief description of the data.

III.A Data Description.

Our dataset consists of 450 (percentage) monthly returns of the 10 stock portfolios formed by size by French, for the period from July 1963 to December 2000[11]. The portfolios are constructed at the end of each June, using the June market equity and NYSE breakpoints. The portfolios for July of year t to June of $t + 1$ include all NYSE, AMEX, and NASDAQ stocks for which we have market equity data for June of year t. Portfolios are ranked by size, with portfolio 1 the smallest, and portfolio 10 the largest.

The market return is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks. The risk free rate is the one-month Treasury bill rate from Ibbotson Associates[12].

III.B Results.

III.B.1 Quadratic Market Model

We begin with the estimation of the quadratic market model (1). PML-SUR estimates of the coefficients α, β, γ and of the variance Σ in model (1) are reported in Tables I and II, respectively.

[**Insert somewhere here Tables I and II**]

As explained in Section II, these estimates are obtained by OLS regressions, performed equation by equation on the system (1). As expected, the beta coefficients are strongly significant for all portfolios, with smaller portfolios having larger betas in general. From the estimates of the γ parameter, we see that small portfolios have significantly negative market coskewness coefficients (for instance $\gamma = -0.017$ for the smallest portfolio), whereas the latter are significantly positive for the two largest portfolios ($\gamma = 0.003$ for the largest portfolio). In particular, we notice that the β and γ coefficients are strongly correlated across portfolios. We can test for joint significancy of the coskewness parameter γ by using the Wald statistics ξ_T^{F*} in (14). It assumes the value: $\xi_T^{F*} = 35.34$, which is strongly significant at the 5 percent level, being the associated critical value $\chi^2_{0.05}(10) = 18.31$. Finally, from Table II, we also see that smaller portfolios are characterized by larger variances of the residual error terms.

We performed several tests of correct specification for the functional form of the mean of portfolios returns in (1). First, we estimated a factor SUR model including also a cubic power of market returns, $R^3_{M,t} - R_{F,t}$, as a factor in addition to the constant, market excess returns and market squared excess returns. The cubic factor is found to be not significant for all portfolios. Furthermore, in order to test for more general forms of misspecifications in the mean, we performed Ramsey Reset Test [Ramsey (1969)] portfolio by portfolio, by including quadratic and cubic fitted values of (1) among the regressors. In this case too, the null of correct specification of the quadratic market model is accepted for all portfolios.

For our analysis, one central result from Table I is that the coskewness coefficients are (significantly) different from zero for all portfolios in our sample, except for two of moderate size. Furthermore, coskewness coefficients tend to be correlated with size, with small portfolios having negative coskewness with the market, and the largest portfolios having positive market coskewness. This result is consistent with the findings of Harvey and Siddique (2000). It is worth noticing that the dependence between portfolios returns and market returns deviates from that of a linear specification (as that assumed in the market model), in directions of smaller (larger) returns for small (large) portfolios when the market has a large absolute return. This finding has important consequences for the assessment of risk in various portfolio classes: small portfolios, having negative market coskewness, are exposed to a source of risk additional to market risk, and related to large absolute market returns. In addition, as we have already seen, the market model (2), if tested against the quadratic market model (1), is rejected with a largely significant Wald statistics. In the light of these findings, we conclude that the extension of the return generating process to include the squared market return is valuable.

III.B.2 Restricted equilibrium models

Let us now investigate market coskewness in the context of models consistent with arbitrage pricing. This is done by considering constrained PML estimation of specification (6), obtained from the quadratic market model after imposing restrictions from the asset pricing model (4), and of specification (8), where a homogeneous additional constant in expected excess returns is allowed for. These PML estimators are obtained from the algorithm based on equations (16) to (18), as reported in Section II [13]. The results for model (6) are reported in Table III and for model (8) in Table IV.

[**Insert somewhere here Tables III and IV**]

The point estimates and standard errors of the parameters β and γ are similar in the two models, and close to those obtained from (1). In particular, the estimates of the parameter γ confirm that small (large) portfolios have significantly negative (positive) coskewness coefficients. The parameter ϑ is found significantly negative in both models, as expected, but the implied estimate for the risk premium for coskewness, $\widehat{\lambda}_2$, is not statistically significant in both models. However, the estimate in model (8), $\widehat{\lambda}_2 = -7.439$, has at least the expected negative sign. Using it, we deduce that, for a portfolio with coskewness $\gamma = -0.01$ (a moderately small portfolio, such as portfolio number 3 or 4), its contribution to the expected excess return on a annual percentage basis is approximately 0.9. This contribution raises to 1.5 for the smallest portfolio in our data set.

We test the empirical validity of the asset pricing model (4) in our sample by testing hypothesis $\mathcal{H}_1$ against the alternative $\mathcal{H}_F$. The ALS test statistics ξ_T^1 given in (21) assumes the value $\xi_T^1 = 16.27$, which is not significant at the 5 percent level, even though very close to the critical value $\chi^2_{0.05}(9) = 16.90$. Thus, there is a modest evidence that asset pricing model (4) could not be satisfied in our sample. In other words, an additional component, other than covariance and coskewness to market, could be present in expected excess returns. In order to test for the homogeneity of this component across assets, we test $\mathcal{H}_2$ against $\mathcal{H}_F$. The test statistics ξ_T^2 in (22) assumes the value of $\xi_T^2 = 5.32$, largely below the critical value $\chi^2_{0.05}(8) = 15.51$. A more powerful test of the asset pricing model (4) should be provided by testing $\mathcal{H}_1$ against the alternative $\mathcal{H}_2$. This test is performed by the simple t-test of significancy of λ_0 and, from Table IV, we see that $\mathcal{H}_1$ is quite clearly rejected. This confirms our finding that asset pricing model (4) may be not supported by the data. However, since $\mathcal{H}_2$ is not rejected, this implies that, if the additional component unexplained by (4) comes from an omitted factor, at least its sensitivities are homogeneous across portfolios in our sample. Thus characteristics such as size and book to market value should not have explanatory power for expected excess returns, when coskewness is taken into account. Moreover, the contribution to expected excess returns of the unexplained component, deduced from the estimate of parameter λ_0, is quite modest, approximately 0.4 on a annual percentage basis. Notice in particular that this is less than the half of the contribution due to coskewness for portfolios of modest size. As explained in section i, $\lambda_0 > 0$ may be due to the use of a risk-free rate lower than the actual rate investors face.

III.B.3 Misspecification from neglected coskewness

As already mentioned in Section I, we are also interested in evaluating the consequences on asset pricing tests of erroneously neglecting coskewness. The results presented so far suggest that the market model (2) is misspecified, given that it does not take into account quadratic market returns. If tested against the quadratic market model (1), it is strongly rejected. For comparison, we report the estimates of the parameter α and β in the market model (2) in Table V.

[**Insert somewhere here Table V**]

We notice that the β coefficients are close to those obtained in the quadratic market model in Table I. Therefore, neglecting the quadratic market returns does not seem to have dramatic consequences for the estimation of parameter β. However, we expect that the consequences of this misspecification to be serious for inference. Indeed, we have seen above that the coskewness coefficients are correlated with size, small portfolios having negative market coskewness and large portfolios positive market coskewness. This suggests that size can have spurious explanatory power in the cross-section of asset expected excess returns since it proxies for omitted coskewness. Therefore, as anticipated in Section I, the empirically observed ability of size to explain expected excess returns could be due to misspecification of models neglecting coskewness risk.

Finally, it is interesting to compare the findings of this paper with those reported in Barone Adesi (1985), whose investigation covers the period 1931-1975. We see that the sign of the premium for coskewness has not changed, with assets having negative coskewness commanding higher expected returns, as expected. On the contrary, both the sign of the premium for size and consequently the link between coskewness and size are inverted. While it appears difficult to discriminate statistically between a structural size effect and reward for coskewness, Kan and Zhang (1999a,b) suggest that persistence of sign and size of coefficients over time are most likely to be associated with truly priced factors. Following them, the explanation of the size effect as arising from neglected coskewness seems to be favored.

IV Monte Carlo simulations.

In this final section, we report the results of a series of Monte Carlo simulations undertaken to investigate the importance of specifying the returns generating process to obtain statistical tests of reliable power. We compare the finite sample properties (size and power) of two statistics for testing the asset pricing model with coskewness (4): i) the ALS statistics ξ_T^1 in (21), which test (4) by the restrictions imposed on the return generating process (1), and ii) a GMM test statistics ξ_T^{GMM}, which tests (4) through the orthogonality conditions (9). In addition, we investigate the effects of nonnormality of the errors ε_t and of model misspecifications of the return generating process (1) on the ALS statistics ξ_T^1.

IV.A Experiment 1.

The data generating process used in Experiment 1 is given by:

$$r_t = \alpha + \beta r_{M,t} + \gamma q_{M,t} + \varepsilon_t,\ t = 1, ..., 450,$$

where $r_{M,t} = R_{M,t} - r_{f,t}$, $q_{M,t} = R_{M,t}^2 - r_{f,t}$, with

$$\begin{aligned} R_{M,t} &\sim iidN(\mu_M, \sigma_M^2), \\ \varepsilon_t &\sim iidN(0, \Sigma),\ (\varepsilon_t) \text{ independent of } (R_{m,t}), \\ r_{f,t} &= r_f, \text{ a constant,} \end{aligned}$$

and

$$\alpha = \vartheta\gamma + \lambda_0 \iota.$$

The values of the parameters are chosen to be equal to the estimates obtained in the empirical analysis reported in the previous section. Specifically, β and γ are the third and fourth columns respectively in Table I, the matrix Σ is taken from Table II, $\vartheta = -14.995$ from Table III, $\mu_m = 0.52$, $\sigma_m = 4.41$, and $r_f = 0.4$, corresponding to the average of the risk free return in our data set. Different values of parameter λ_0 are used in the simulations. We will refer to this data generating process as DGP1. Under DGP1, when $\lambda_0 = 0$, the quadratic equilibrium model (4) is satisfied. When $\lambda_0 \neq 0$, the equilibrium model (4) is not correctly specified, and the misspecification is in the form of an additional component homogeneous across portfolios, corresponding to model (8). However, the quadratic model (1) is in any case well-specified.

We perform Monte Carlo simulation (10000 replications), for different values of λ_0, and report the rejection frequencies of the two test statistics, ξ_T^1 and ξ_T^{GMM}, at the nominal size of 0.05 in Table VI.

[Insert somewhere here Table VI]

The second row, $\lambda_0 = 0$, reports the empirical sizes. Both statistics control the size quite well in finite sample, at least for sample size $T = 450$. The subsequent rows, corresponding to $\lambda_0 \neq 0$, report the power of the two test statistics against alternatives corresponding to unexplained components in expected excess returns, which are homogeneous across portfolios. Note that such additional components, with $\lambda_0 = 0.033$, were found in the data in the empirical analysis. Table VI shows that the power of the ALS statistics ξ_T^1 is considerably higher than that of the GMM statistics ξ_T^{GMM}. This is due to the fact that the ALS statistics ξ_T^1 uses a well-specified alternative for testing (1), whereas the alternative for the GMM statistics ξ_T^{GMM} is left unspecified.

IV.B Experiment 2.

Under DGP1, the residuals ε_t are normal. When the ε_t are not normal, the alternative used by the ALS statistics ξ_T^1, that is model (1), is still correctly specified, since PML estimators are used to construct ξ_T^1. However, these estimators are not efficient. In experiment 2 we investigate the effect on the ALS test statistics of non-normality of the residuals ε_t. The data generating process used in this experiment, called DGP2, is equal to DGP1 but the residuals ε_t follow a multivariate t-distribution with $df = 5$ degrees of freedom, and a correlation matrix such that variance of ε_t is the same as under DGP1. The rejection frequencies of the Monte Carlo simulation (10000 replications) for the ALS statistics ξ_T^1 are reported in Table VII.

[**Insert somewhere here Table VII**]

The ALS statistics appears to be only slightly oversized. As expected, the power is reduced compared to the case of normality, however the loss of power caused by non-normality is limited. These results suggest that the ALS statistics does not unduly suffer from departures from normality of the residuals.

IV.C Experiment 3.

In the experiments conducted so far, the alternative used by the ALS statistics was well-specified. In this last experiment we investigate the effect of a misspecification in the alternative in the form of conditional heteroscedasticity. We thus consider two data generating processes having the same unconditional variance of the residuals ε_t, but such that the residuals ε_t are conditionally heteroscedastic in one case, and homoscedatic in the other. Specifically, DGP3 is the same as DGP1, but the innovations ε_t follow a conditionally normal, multivariate ARCH(1) process without cross effects:

$$cov\left(\varepsilon_{i,t}, \varepsilon_{j,t} \mid \underline{\varepsilon_{t-1}}\right) = \begin{cases} \omega_{ii} + \rho\varepsilon_{i,t-1}^2, & i = j \\ \omega_{ij}, & i \neq j \end{cases} .$$

The matrix $= [\omega_{ij}]$ is chosen as in Table II, and $\rho = 0.2$. DGP4 is the same as DGP1, with i.i.d. normal innovations whose unconditional variance matrix is the same as the unconditional variance of ε_t in DGP 3. Thus under DGP4 the alternative of the ALS statistics is well-specified, but not under

DGP3. The rejection frequencies of the ALS statistics under DGP3 and DGP4 are reported in Table VIII.

[**Insert somewhere here Table VIII**]

The misspecification in form of conditional heteroscedasticity has no effect on the empirical size of the statistics in these simulations. The power of the test is reduced, but not dramatically.

V Conclusions

In this paper we consider coskewness and its implications for testing asset pricing models. We use a quadratic market model as return generating process, with market returns and the square of market returns as the two factors. It is shown that portfolios of small (large) firms have negative (positive) coskewness with market. This implies that small portfolios are subject to a further source of risk other than covariance with market, that is market coskewness, which arises from (negative) covariance with large absolute market returns. Coskewness coefficients of the portfolios in our sample are shown to be jointly significant, rejecting the usual market model. These findings imply that the quadratic market model, used as a return generating process, is a valuable extension of the market model.

In order to obtain methodologies of superior power, we propose to test an asset pricing model, including coskewness, through the restrictions it imposes on the return generating process. We use an asymptotic test statistics whose finite sample properties are validated via a series of Monte Carlo simulations. We find evidence for a component in expected excess returns which is not explained by neither covariance nor coskewness with the market. We show that this unexplained component in expected excess returns does not vary across portfolios and it is modest in magnitude. This is consistent with a minor misspecification of the risk-free rate. This finding implies that additional variables representing portfolios characteristics we consider have no explanatory power for expected excess returns when coskewness is taken in account. This result cannot be obtained if coskewness is neglected.

In addition to that, our results have implications for testing methodologies, since they show that neglecting coskewness risk can cause misleading inference. Indeed, we find that coskewness is positively correlated with size. This suggests that a possible justification for the anomalous explanatory power of size in the cross-section of expected returns, is that it proxies for omitted coskewness risk. This view is supported by the fact that the sign of the premium for coskewness, contrary to that of size, has not changed over time.

APPENDICES

Appendix A: PML in model (8).

In this appendix we consider the Pseudo Maximum Likelihood (PML) estimator of model (8), defined by maximization of (15). Let us first derive the PML equations. The score vector is given by:

$$
\begin{aligned}
\frac{\partial L_T}{\partial \left(\beta', \gamma'\right)'} &= \sum_{t=1}^{T} H_t \quad \Sigma^{-1}\varepsilon_t, \\
\frac{\partial L_T}{\partial \left(\vartheta, \lambda_0\right)'} &= \sum_{t=1}^{T} Z'\Sigma^{-1}\varepsilon_t, \\
\frac{\partial L_T}{\partial vech(\Sigma)} &= \frac{1}{2}P^T\Sigma^{-1} \quad \Sigma^{-1}Pvech\left[\sum_{t=1}^{T}\left(\varepsilon_t\varepsilon_t' - \Sigma\right)\right],
\end{aligned}
$$

where $H_t = \left(r_{M,t}, q_{M,t} + \vartheta\right)'$, $\varepsilon_t = r_t - \beta r_{M,t} - \gamma q_{M,t} - \gamma\vartheta - \lambda_0\iota$, $Z = (\gamma, \iota)$ and P is such that $vec\,(\Sigma) = Pvech\,(\Sigma)$. By equating the score to 0, we immediately find the equations (16) to (18).

Let us now derive the asymptotic distribution of the PML estimator in model (8). Under usual regularity conditions (see references in the test) the asymptotic distribution of the general PML estimator $\widehat{\theta}$ defined in (11) is given by:

$$
\sqrt{T}\left(\widehat{\theta} - \theta^0\right) \xrightarrow{d} N\left(0, J_0^{-1} I_0 J_0^{-1}\right),
$$

where J_0 (the so called information matrix), and I_0 are symmetric, positive definite matrices defined by:

$$
J_0 = \lim_{T\to\infty} E\left[-\frac{1}{T}\frac{\partial^2 L_T}{\partial\theta\partial\theta'}(\theta^0)\right], \quad I_0 = \lim_{T\to\infty} E\left[\frac{1}{T}\frac{\partial L_T}{\partial\theta}(\theta^0)\frac{\partial L_T}{\partial\theta'}(\theta^0)\right].
$$

Let us compute matrices J_0 and I_0 in model (8). The second derivatives of the loglikelihood are given by:

$$
\begin{aligned}
\frac{\partial^2 L_T}{\partial\left(\beta^{'},\gamma^{'}\right)^{'}\partial\left(\beta^{'},\gamma^{'}\right)} &= -\sum_{t=1}^{T} H_t H_t^{'} \quad \Sigma^{-1}, \\
\frac{\partial^2 L_T}{\partial\left(\beta^{'},\gamma^{'}\right)^{'}\partial\left(\vartheta,\lambda_0\right)^{'}} &= -\sum_{t=1}^{T} H_t \quad \Sigma^{-1} Z, \\
\frac{\partial^2 L_T}{\partial\left(\vartheta,\lambda_0\right)\partial\left(\vartheta,\lambda_0\right)^{'}} &= -T Z^{'}\Sigma^{-1} Z, \\
\frac{\partial L_T}{\partial vech(\Sigma)\partial vech(\Sigma)^{'}} &= \frac{T}{2} P^T \Sigma^{-1} \quad \Sigma^{-1} P - \frac{T}{2} P^T \Sigma^{-1} \quad \Sigma^{-1}\left(\sum_{t=1}^{T}\varepsilon_t\varepsilon_t^{'}\right)\Sigma^{-1} P \\
&\quad -\frac{T}{2} P^T \Sigma^{-1}\left(\sum_{t=1}^{T}\varepsilon_t\varepsilon_t^{'}\right)\Sigma^{-1} \quad \Sigma^{-1} P,
\end{aligned}
$$

with the other ones vanishing in expectation. It results that matrices J_0 and I_0 are given by [in the block representation corresponding to $\left(\beta^{'},\gamma^{'},\vartheta,\lambda_0\right)^{'}$ and $vech\,(\Sigma)$]:

$$
J_0 = \begin{bmatrix} J_0^* & \\ & \widetilde{J}_0 \end{bmatrix}, \qquad I_0 = \begin{bmatrix} J_0^* & \eta S \widetilde{J}_0 \\ \widetilde{J}_0 S^{'} \eta^{'} & \widetilde{J}_0 K \widetilde{J}_0 \end{bmatrix},
$$

where:

$$
\widetilde{J}_0 = \frac{1}{2}\left(P^T \Sigma^{-1} \quad \Sigma^{-1} P\right),
$$

$$
S = cov\left[\varepsilon_t, vech\left(\varepsilon_t\varepsilon_t^{'}\right)^{'}\right], \quad K = Var\left[vech\left(\varepsilon_t\varepsilon_t^{'}\right)\right],
$$

and, in the block form corresponding to $\left(\beta^{'},\gamma^{'}\right)^{'}$, $(\vartheta,\lambda_0)^{'}$:

$$
J_0^* = \begin{bmatrix} E\left[H_t H_t\right] \quad \Sigma^{-1} & \lambda \quad \Sigma^{-1} Z \\ \lambda^{'} \quad Z^{'}\Sigma^{-1} & Z^{'}\Sigma^{-1} Z \end{bmatrix}, \eta = \begin{bmatrix} \lambda \quad \Sigma^{-1} \\ Z^{'}\Sigma^{-1} \end{bmatrix}.
$$

(All parameters are evaluated at the true value). Therefore, the asymptotic variance-covariance matrix of the PML estimator $\widehat{\theta}$ in model (8) is given by:

$$
V_{as}\left[\sqrt{T}\left(\widehat{\theta}-\theta_0\right)\right] = J_0^{-1} I_0 J_0^{-1} = \begin{bmatrix} J_0^{*-1} & J_0^{*-1}\eta S \\ S^{'}\eta^{'} J_0^{*-1} & K \end{bmatrix}.
$$

Notice that the asymptotic variance-covariance of $\left(\widehat{\beta}^{'}, \widehat{\gamma}^{'}, \widehat{\vartheta}, \widehat{\lambda}_0\right)^{'}$, that is J_0^{*-1}, does not depend on the distribution of the error term ε_t, and in particular it coincides with the asymptotic variance-covariance matrix of the maximum likelihood (ML) estimator of $\left(\widehat{\beta}^{'}, \widehat{\gamma}^{'}, \widehat{\vartheta}, \widehat{\lambda}_0\right)^{'}$ when ε_t is normal. On the contrary, asymmetries and kurtosis of the distribution of ε_t influence the asymptotic variance-covariance matrix of $vech(\Sigma)$ and the asymptotic covariance of $\left(\widehat{\beta}^{'}, \widehat{\gamma}^{'}, \widehat{\vartheta}, \widehat{\lambda}_0\right)^{'}$ and $vech(\widehat{\Sigma})$, through matrices S and K.

The asymptotic variance-covariance of $\left(\widehat{\beta}^{'}, \widehat{\gamma}^{'}\right)^{'}$ and $\left(\widehat{\vartheta}, \widehat{\lambda}_0\right)^{'}$ is given explicitly in block form by:

$$J_0^{*-1} = \begin{bmatrix} J_0^{*11} & J_0^{*12} \\ J_0^{*21} & J_0^{*22} \end{bmatrix},$$

where:

$$J_0^{*11} = \left(\Sigma_f + \lambda\lambda^{'}\right)^{-1} \quad \Sigma + \left[\Sigma_f^{-1}\lambda\lambda^{'}\left(\Sigma_f + \lambda\lambda^{'}\right)^{-1}\right] \quad Z\left(Z^{'}\Sigma^{-1}Z\right)^{-1} Z^{'},$$

$$J_0^{*12} = -\Sigma_f^{-1}\lambda \quad Z\left(Z^{'}\Sigma^{-1}Z\right)^{-1},$$

$$J_0^{*21} = J_0^{*12'}$$

$$J_0^{*22} = \left(1 + \lambda^{'}\Sigma_f^{-1}\lambda\right)\left(Z^{'}\Sigma^{-1}Z\right)^{-1}.$$

Finally, let us consider the asymptotic distribution of estimator $\widehat{\lambda}$ defined in (19). The estimator:

$$\widehat{\mu} = \frac{1}{T}\sum_{t=1}^{T} f_t,$$

where $f_t = (r_{M,t}, q_{M,t})^{'}$, can be seen as a component of the PML estimator on the extended pseudo-likelihood:

$$L_T(\theta, \mu, \Sigma_f) = L_T(\theta) - \frac{T}{2}\log\det\Sigma_f - \frac{1}{2}\sum_{t=1}^{T}(f_t - \mu)^{'}\Sigma_f^{-1}(f_t - \mu),$$

where $L_T(\theta)$ is given in (15). It is easily seen that θ and (μ, Σ_f) are asymptotically independent. It follows:

$$V_{as}\left[\sqrt{T}\left(\widehat{\lambda}_2 - \lambda_{2,0}\right)\right] = \Sigma_{f,22} + V_{as}\left[\sqrt{T}\left(\widehat{\vartheta} - \vartheta_0\right)\right].$$

Appendix B: Asymptotic Least Squares.

In this Appendix we derive the ALS statistics ξ_T^1 in (21) and ξ_T^2 in (22). In both cases the restrictions [see (20)] are of the form:

$$g(\theta, a) = A_1(a)vec(B) + A_2(a),$$

where B is the $N \times 3$ matrix defined by $B = [\alpha\ \beta\ \gamma]$ and $A_1(a)$ is such that:

$$A_1(a) = (1, 0, -\vartheta) \quad I_N \equiv A_1^*(a) \quad I_N.$$

Let us derive the weighting matrix $S_0 = (\partial g/\partial \theta^{'} \quad {}_0 \partial g/\partial \theta)^{-1}$, where $\quad {}_0 = V_{as}\left(\sqrt{T}\left(\widehat{\theta} - \theta\right)\right)$. >From (13) we get:

$$\begin{aligned} \frac{\partial g}{\partial \theta^{'}} \quad {}_0 \frac{\partial g^{'}}{\partial \theta} &= A_1^* E\left[F_t F_t^{'}\right]^{-1} A_1^{*'} \quad \Sigma \\ &= \left(1 + \lambda^{'} \Sigma_f^{-1} \lambda\right) \Sigma. \end{aligned}$$

The test statistics follow.

REFERENCES

Banz, R. W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

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

Barone Adesi, G., and P. Talwar, 1983, Market models and heteroscedasticity of residual security returns, *Journal of Business and Economic Statistics 1,* 163-168.

Black, F., 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444-55.

Bollerslev, T., and J. M. Wooldridge, 1992, Quasi Maximum Likelihood estimation and inference in dynamic models with time varying covariances, *Econometric Reviews* 11, 143-172.

Breeden, D. T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265-96.

Campbell, J. Y., Lo, A. W., and A. C. MacKinlay, 1987. *The Econometrics of Financial Markets* (Princeton University Press, New Jersey).

Chamberlain, G., and M. Rothschild, 1983, Arbitrage, factor structure and mean variance analysis in large asset markets, *Econometrica* 51, 1281-1301.

Dittmar, R. F., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from cross section of equity returns, *Journal of Finance* 57, 368-403.

Engle, R. F., Hendry, D. F., and J. F. Richard, 1988, Exogeneity, *Econometrica* 51, 277-304.

Fama, E., and K. R. French, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131-55.

Gibbons, M., S. Ross, and J. Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Gourieroux, C., and J. Jasiak, 2001. *Financial Econometrics* (Princeton University Press).

Gourieroux, C., Monfort, A., and A. Trognon, 1984, Pseudo Maximum Likelihood methods: Theory, *Econometrica* 52, 681-700.

Gourieroux, C., Monfort, A., and A. Trognon, 1985, Moindres Carrés Asymptotiques, *Annales de l'INSEE* 58, 91-122.

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica 50, 1029-1054.*

Hansen, L. P., and R. Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557-590.

Harvey, C.R., and A. Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.

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

Jagannathan, R., and Z. Wang, 2001, Empirical evaluation of asset pricing models: A comparison of the SDF and beta methods", Working Paper, Northwestern University, Illinois.

Kan, R., and C. Zhang, 1999a, GMM test of stochastic discount factor models with useless factors, *Journal of Financial Economics* 54, 103-27.

Kan, R., and C. Zhang, 1999b, Two-pass tests of asset pricing models with useless factors, *Journal of Finance* 54, 203-35.

Kan, R., and G. Zhou, 1999, A critique of the stochastic discount factor methodology, *Journal of Finance* 54, 1221-1248.

Kraus, A., and R. Litzenberger, 1976, Skewness preferences and the valuation of risk assets, *Journal of Finance* 31, 1085-1100.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

MacKinlay, A.C., and L. Pastor, 2000, Asset pricing models: Implications for expected returns and portfolio selection, *The Review of Financial Studies* 13, 883-916.

Merton, R. C., 1973, An intertemporal capital asset pricing model," *Econometrica* 41, 867-87.

Newey, W. K., and K. D. West, 1989, A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-8.

Ramsey, J. B., 1969, Tests for specification errors in classical linear least squares regression analysis, *Journal of the Royal Statistical Society, Series B,* 31, 350-371.

Ross, S. A., 1976, Arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-60.

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

Sharpe, W. F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk , *Journal of Finance* 40, 1189-96.

Velu, R., and G. Zhou, 1999, Testing multi-beta asset pricing models, *Journal Of Empirical Finance* (6)3, 219-241.

White, H., 1981, Maximum Likelihood estimation of misspecified models, *Econometrica* 50, 1-25.

Zellner, A., 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of American Statistical Association* 57, 348-68.

Zhou, G., 1995, Small sample rank tests with applications to asset pricing, *Journal of Empirical Finance* (2)1, 71-93.

Notes

[1]Coskewness cannot be explained e.g. in the framework of MacKinlay and Pastor (2000), because their assumption of multivariate normality implies independence of the unexplained returns to the tested factor.

[2]Our methodology is similar in spirit to one of the approaches of Harvey and Siddique (2000), who include a coskewness portfolio among the regressors in an extended market model, and test the corresponding asset pricing model by a Gibbons-Ross-Shanken F-statistics [Gibbons, Ross, Shanken (1989)]. In our model, testing methodologies are less simple, since the square of the market returns is not a traded asset.

[3]Harvey and Siddique (2000), p. 1281, suggest that "...book/market value and size effects in asset returns may proxy for conditional skewness in asset returns".

[4]As another consequence on inference of neglecting erroneously coskewness, we expect the power of the return generating process to be seriously compromised. As an example, we can compare results for testing the Capital Asset Pricing Model (CAPM) when the market model is the alternative, and when the quadratic market model is the alternative. If the market model is misspecified due to the omission of the quadratic market returns term, its power is presumably low against CAPM.

[5]for a time series $(Y_t, t \in \mathbb{Z})$, $\underline{Y_t}$ denotes all present and past values Y_s, $s \leq t$.

[6]Note that γ can equivalently be written as:

$$\gamma = \frac{1}{V\left[\epsilon_{q,t}\right]} cov\left[\epsilon_t, \epsilon_{q,t}\right].$$

The numerator is a third-order cross moment of the residuals in the regressions of R_t and $R_{m,t}^2$ on $R_{m,t}$. This is slightly different from the measure of coskewness of Kraus and Litzenberger (1976)

[7]Ross (1976) shows that the absence of arbitrage implies the approximative relation: $E(r_t) \simeq \beta\lambda_1 + \gamma\lambda_2$. Under additional restrictions this relation becomes exact [see e.g. the discussion in Campbell, Lo, MacKinlay (1987)]. In this paper we assume as usual that exact factor pricing holds.

[8]For a $n \times n$ symmetric matrix A, $vech(A)$ denotes the $\frac{(n+1)n}{2} \times 1$ vector representation of A, where only elements on and above the main diagonal appear.

[9]Upper indexes in a matrix denote elements of the inverse.

[10]It should be noted that exact tests (under normality) can be constructed for testing hypotheses $\mathcal{H}_1$ and $\mathcal{H}_2$ against $\mathcal{H}_F$ [see e.g. Zhou (1995), and Velu and Zhou (1999)]. These tests are asymptotically equivalent to the Asymptotic Least Squares tests, which are proposed in the paper for their computational simplicity. A first assessement of the finite sample properties of the ALS test statistics is presented in section IV.

[11]Data are available from the site http://web.mit.edu/kfrench/www/data_library.html, in the file "Portfolios Formed on Size".

[12]The market return and risk free return are available from the site http://web.mit.edu/kfrench/www/data_library.html, in the files "Fama-French Benchmark Factors" and "Fama-French Factors". We use the T-bill rate because other money-market series are not available for the whole period of our tests.

[13]As convergence criterium we required the update of each parameter to be smaller than $5 \cdot 10^{-3}$. We obtained convergence of the algorithm to the same estimates over the range of sensible alternative starting points we have tried.

Table I: Coefficient estimates of model (1).

Portfolio i	$\widehat{\alpha}_i$	$\widehat{\beta}_i$	$\widehat{\gamma}_i$
1	0.418 (1.84) [1.70]	1.101 (24.23) [20.24]	−0.017 (−3.32) [−2.94]
2	0.299 (1.65) [1.56]	1.188 (32.62) [27.07]	−0.013 (−3.05) [−2.65]
3	0.288 (1.88) [1.86]	1.182 (38.37) [29.18]	−0.010 (−2.84) [−2.45]
4	0.283 (1.96) [1.83]	1.166 (39.99) [30.98]	−0.010 (−3.00) [−2.82]
5	0.328 (2.73) [2.51]	1.135 (46.94) [34.16]	−0.009 (−3.34) [−2.68]
6	0.162 (1.59) [1.53]	1.110 (54.02) [37.85]	−0.006 (−2.58) [−2.28]
7	0.110 (1.29) [1.24]	1.105 (64.37) [50.66]	−0.002 (−0.88) [−0.84]
8	0.076 (1.02) [0.90]	1.083 (72.59) [56.61]	−0.000 (−0.18) [−0.23]
9	−0.016 (−0.30) [−0.28]	1.017 (92.76) [98.43]	0.003 (2.06) [2.26]
10	−0.057 (−1.10) [−0.99]	0.933 (88.77) [66.71]	0.003 (2.64) [2.73]

Notes: Table I reports for each portfolio i, $i = 1, ..., 10$, the PML-SUR estimates of the coefficients α_i, β_i, γ_i of the quadratic market model:

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \gamma_i q_{M,t} + \varepsilon_{i,t},\ t = 1, ..., T,\ \ i = 1, ..., N,$$

where $r_{i,t} = R_{i,t} - R_{F,t}$, $r_{M,t} = R_{M,t} - R_{F,t}$, $q_{M,t} = R^2_{M,t} - R_{F,t}$. $R_{i,t}$ is the return of portfolio i in month t, and $R_{M,t}$ ($R_{F,t}$) denotes the market return (the risk free return). In round parentheses we report *t-statistics* computed under the assumption:

$$\begin{aligned} E\left[\varepsilon_t | \underline{\varepsilon_{t-1}}, \underline{R_{M,t}}, \underline{R_{F,t}}\right] &= 0, \\ E\left[\varepsilon_t \varepsilon_t^{'} | \underline{\varepsilon_{t-1}}, \underline{R_{M,t}}, \underline{R_{F,t}}\right] &= \Sigma, \qquad \varepsilon_t = \left(\varepsilon_{1,t}, ..., \varepsilon_{N,t}\right), \end{aligned}$$

while *t-statistics* calculated with Newey-West (1987) heteroscedasticity and autocorrelation consistent estimator with 5 lags are in square parentheses.

Table II: Variance estimates of model (1).

	1	2	3	4	5	6	7	8	9	10
1	17.94	13.42	10.69	9.41	6.93	5.20	4.02	2.64	0.51	−3.11
2		11.50	9.02	8.27	6.35	4.81	3.69	2.61	0.58	−2.72
3			8.24	7.18	5.65	4.51	3.34	2.39	0.68	−2.40
4				7.39	5.56	4.37	3.40	2.41	0.78	−2.33
5					5.07	3.71	2.82	2.21	0.77	−1.93
6						3.67	2.42	1.85	0.78	−1.59
7							2.56	1.68	0.75	−1.29
8								1.93	0.85	−1.05
9									1.04	−0.50
10										0.96

Notes: Table II reports the estimate of the variance $\Sigma = E\left[\varepsilon_t \varepsilon_t' \mid \underline{r_{M,t}}, \underline{q_{M,t}}\right]$ of the error ε_t in the quadratic market model:

$$r_t = \alpha + \beta r_{M,t} + \gamma q_{M,t} + \varepsilon_t,\ t = 1, ..., T,\ \ i = 1, ..., N,$$

where $r_t = R_t - R_{F,t}\iota$, $r_{M,t} = R_{M,t} - R_{F,t}$, $q_{M,t} = R_{M,t}^2 - R_{F,t}$. R_t is the N-vector of portfolios returns, $R_{M,t}$ ($R_{F,t}$) is the market return (the risk free return), and ι is a N-vector of ones.

Table III: PML estimates of model (6).

Portfolio i	$\widehat{\beta}_i$	$\widehat{\gamma}_i$
1	1.106 (24.50)	−0.017 (−3.25)
2	1.191 (32.97)	−0.012 (−2,99)
3	1.186 (38.79)	−0.009 (−2.74)
4	1.170 (40.41)	−0.009 (−2.90)
5	1.140 (47.38)	−0.009 (−3.14)
6	1.112 (54.56)	−0.006 (−2.50)
7	1.107 (65.07)	−0.001 (−0.76)
8	1.085 (73.37)	−0.001 (−0.05)
9	1.017 (93.66)	0.002 (2.14)
10	0.933 (89.53)	0.003 (2.63)

$\widehat{\vartheta} = -14.955$ (−2.23)	$\widehat{\lambda}_2 = 4.850$ (0.70)

Notes: Table III reports PML estimates of the coefficients of the restricted model (6):

$$r_t = \beta r_{M,t} + \gamma q_{M,t} + \gamma\vartheta + \varepsilon_t, t = 1, ..., T,$$

where ϑ is a scalar parameter, derived from the quadratic market model (1) by imposing the restriction given by the asset pricing model with coskewness: $E(r_t) = \lambda_1\beta + \lambda_2\gamma$. The scalar ϑ and the premium for coskewness λ_2 are related by: $\vartheta = \lambda_2 - E(q_{M,t})$. The restricted model (6) corresponds to hypothesis $\mathcal{H}_1$: $\exists\vartheta : \alpha = \vartheta\gamma$ in (1). *t-statistics* are reported in parentheses.

Table IV: PML estimates of model (8).

Portfolio i	$\widehat{\beta}_i$	$\widehat{\gamma}_i$
1	1.100 (24.38)	−0.017 (−3.32)
2	1.187 (32.84)	−0.012 (−3.05)
3	1.183 (38.70)	−0.010 (−2.91)
4	1.167 (40.31)	−0.010 (−3.07)
5	1.137 (47.35)	−0.009 (−3.52)
6	1.110 (54.45)	−0.006 (−2.62)
7	1.107 (65.07)	−0.002 (−1.06)
8	1.085 (73.40)	−0.001 (−0.38)
9	1.018 (93.72)	0.002 (1.90)
10	0.934 (89.60)	0.003 (2.57)

$\widehat{\vartheta} = -27.244$ (−3.73)	$\widehat{\lambda}_2 = -7.439$ (1.01)	$\widehat{\lambda}_0 = 0.032$ (3.27)

Notes: Table IV reports PML estimates of the coefficients of the restricted model (8):

$$r_t = \beta r_{M,t} + \gamma q_{M,t} + \gamma\vartheta + \lambda_0 \iota + \varepsilon_t, t = 1, ..., T,$$

where ϑ and λ_0 are scalar parameters, derived from the quadratic market model (1) by imposing the restriction: $E(r_t) = \lambda_0 \iota + \lambda_1 \beta + \lambda_2 \gamma$. Under this restriction, asset expected excess returns contain a component λ_0 which is not explained by neither covariance nor coskewness with the market. The restricted model (8) corresponds to hypothesis $\mathcal{H}_2$: $\exists \vartheta, \lambda_0 : \alpha = \vartheta\gamma + \lambda_0 \iota$ in (1). *t-statistics* are reported in parentheses.

Table V: Estimates of model (2).

Portfolio i	$\widehat{\alpha}_i$	$\widehat{\beta}_i$
1	0.080 (0.39)	1.102 (23.97)
2	0.050 (0.31)	1.188 (32.34)
3	0.092 (0.67)	1.183 (38.09)
4	0.088 (0.67)	1.167 (39.65)
5	0.148 (1.36)	1.135 (46.43)
6	0.044 (0.48)	1.110 (53.69)
7	0.076 (1.00)	1.105 (64.39)
8	0.069 (1.05)	1.083 (72.67)
9	0.034 (0.71)	1.017 (92.41)
10	0.005 (0.10)	0.933 (88.18)

Notes: Table V reports for each portfolio i, $i = 1, ..., 10$, the PML-SUR estimates of the coefficients α_i, β_i of the traditional market model :

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t},\ t = 1, ..., T,\ \ i = 1, ..., N,$$

where $r_{i,t} = R_{i,t} - R_{F,t}$, $r_{M,t} = R_{M,t} - R_{F,t}$. $R_{i,t}$ is the return of portfolio i in month t, and $R_{M,t}$ ($R_{F,t}$) is the market return (the risk free return). *t-statistics* are reported in parentheses.

Table VI: Rejection frequencies in experiment 1

λ_0	ξ_T^{GMM}	ξ_T^1
0.00	0.0404	0.0559
0.03	0.0505	0.4641
0.06	0.0712	0.9746
0.10	0.1217	0.9924
0.15	0.2307	0.9945

Notes: Table VI reports the rejection frequencies of the GMM statistics ξ_T^{GMM} [derived from (9)] and the ALS statistics ξ_T^1 [in (21)] for testing the asset pricing model with coskewness (4):

$$E(r_t) = \lambda_1 \beta + \lambda_2 \gamma,$$

at 0.05 confidence level, in experiment 1. The data generating process (called DGP1) used in this experiment is given by:

$$r_t = \alpha + \beta r_{M,t} + \gamma q_{M,t} + \varepsilon_t,\ t = 1, ..., 450,$$

where $r_{M,t} = R_{M,t} - r_{f,t}$, $q_{M,t} = R_{M,t}^2 - r_{f,t}$, with

$$\begin{aligned} R_{M,t} &\sim iidN(\mu_M, \sigma_M^2), \\ \varepsilon_t &\sim iidN(0, \Sigma),\ (\varepsilon_t) \text{ independent of } (R_{m,t}), \\ r_{f,t} &= r_f, \text{ a constant,} \end{aligned}$$

and

$$\alpha = \vartheta \gamma + \lambda_0 \iota.$$

Parameters β and γ are the third and fourth columns respectively in Table I, the matrix Σ is taken from Table II, $\vartheta = -14.995$ from Table III, $\mu_m = 0.52$, $\sigma_m = 4.41$, and $r_f = 0.4$, corresponding to the average of the risk free return in our data set. Under DGP1, when $\lambda_0 = 0$, the quadratic equilibrium model (4) is satisfied. When $\lambda_0 \neq 0$, the equilibrium model (4) is not correctly specified, and the misspecification is in the form of an additional component homogeneous across portfolios, corresponding to model (8).

Table VII: Rejection frequencies in experiment 2

λ_0	ξ_T^1
0.00	0.0617
0.03	0.3781
0.06	0.9368
0.10	0.9876
0.15	0.9910

Notes: Table VII reports the rejection frequencies of the ALS statistics ξ_T^1 [in (21)] for testing (4):

$$E(r_t) = \lambda_1 \beta + \lambda_2 \gamma,$$

at 0.05 confidence level, in experiment 2. The data generating process used in this experiment (called DGP2) is the same as DGP1 (see Table VI), but the residuals ε_t follow a multivariate t-distribution with $df = 5$ degrees of freedom, and a correlation matrix such that variance of ε_t is the same as under DGP1.

Table VIII: Rejection frequencies in experiment 3

λ_0	ξ_T^1 under DGP 4 (homosced.)	ξ_T^1 under DGP 3 (cond. heterosced.)
0.00	0.0587	0.0539
0.03	0.3683	0.1720
0.06	0.9333	0.5791
0.10	0.9855	0.9373

Notes: Table VIII reports the rejection frequencies of the ALS statistics ξ_T^1 [in (21)] for testing (4):

$$E(r_t) = \lambda_1 \beta + \lambda_2 \gamma,$$

at 0.05 confidence level, in experiment 3. In this experiment we consider two data generating processes (called DGP3 and DGP4) having the same unconditional variance of the residuals ε_t, but such that the residuals ε_t are conditionally heteroscedastic in one case, and homoscedatic in the other. Specifically, DGP3 is the same as DGP1 (see Table VI), but the innovations ε_t follow a conditionally normal, multivariate ARCH(1) process without cross effects:

$$cov\left(\varepsilon_{i,t}, \varepsilon_{j,t} \mid \underline{\varepsilon_{t-1}}\right) = \begin{cases} \omega_{ii} + \rho \varepsilon_{i,t-1}^2, & i = j \\ \omega_{ij}, & i \neq j \end{cases} .$$

The matrix $\quad = [\omega_{ij}]$ is chosen as in Table II, and $\rho = 0.2$. DGP4 is the same as DGP1 (see Table VI), with i.i.d. normal innovations whose unconditional variance matrix is the same as the unconditional variance of ε_t in DGP 3. Thus under DGP4 the alternative of the ALS statistics is well-specified, but not under DGP3.