Testing Asset Pricing Models with Coskewness

Giovanni BARONE ADESI
Facolta' di Finanza, Universita' della Svizzera Italiana, Via Buffi 13, 6900, Lugano (giovanni.barone-adesi@lu.unisi.ch)
Patrick GAGLIARDINI
Facolta' di Finanza, Universita' della Svizzera Italiana, Via Buffi 13, 6900, Lugano (patrick.gagliardini@lu.unisi.ch)
Giovanni URGA
Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ (g.urga@city.ac.uk)

In this paper we investigate portfolio coskewness using a quadratic market model as a return-generating process. We show that the portfolios of small (large) firms have negative (positive) coskewness with the market. An asset pricing model including coskewness is tested by checking the validity of the restrictions it imposes on the return generating process. We find evidence of an additional component in expected excess returns, which is not explained by either covariance or coskewness with the market. However, this unexplained component is homogeneous across portfolios in our sample and modest in magnitude. Finally, we investigate the implications of erroneously neglecting coskewness for testing asset pricing models, with particular attention to the empirically detected explanatory power of firm size.

KEY WORDS: Asset pricing, Coskewness, Asymptotic least squares, Generalized method of moments, Monte Carlo simulations.

1 INTRODUCTION

Asset pricing models generally express expected returns on financial assets as linear functions of covariances of returns with some systematic risk factors. Several formulations of this general paradigm have been proposed in the literature [Sharpe 1964; Lintner 1965; Black 1972; Merton 1973; Rubinstein 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 and 2000; Dittmar 2002]. However, most of the empirical tests suggested to date have produced negative or ambiguous results. These findings have spurred renewed interest in the statistical properties of the currently available testing methodologies. Among recent studies, Shanken (1992) and Kan and Zhang (1999a,b) analyse the statistical methodologies commonly employed and highlight the sources of ambiguity in 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 a-priori knowledge may lead to the design of reliable tests. Since 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 not be powerful, as noted by Kan and Zhou (1999) and Jagannathan and Wang (2001). Note that the first choice may lead not only to the rejection of the correct model, 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, which are not consistent with standard asset pricing models such as the mean-variance Capital Asset Pricing Model (CAPM). Among other studies, Banz (1981) relates expected returns to firm size, while 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 description of market equilibrium. These pricing anomalies may be related to the possibility that useless factors appear to be priced. Of course it is also possible that pricing anomalies are due to omitted factors. While statistical tests do not allow us to choose between these two possible explanations of pricing anomalies, Kan and Zhang (1999a,b) suggest that perhaps a large increment in R^2 and the persistence of sign and size of coefficients over time are most likely to be associated with truly priced factors.

In the light of the above, the aim of this paper is to consider market coskewness and to investigate its role in testing asset pricing models. A data set of monthly returns on 10 stock portfolios is used. Following Harvey and Siddique (2000), an asset is defined as having "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 reduces (increases) the risk of the portfolio to large absolute market returns, and should command a lower (higher) expected return in equilibrium.

Rubinstein (1973), Kraus and Litzenberger (1976), Barone-Adesi (1985) and Harvey and Siddique (2000) study non-normal asset pricing models related to coskewness. Kraus and Litzenberger (1976) and Harvey and Siddique (2000) formulate expected returns as a function of covariance and coskewness with the market portfolio. In particular, Harvey and Siddique (2000) assess the importance of coskewness in explaining expected returns by the increment 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. The author tests an extended asset pricing model within a Generalized Method of Moment (GMM) framework (see Hansen 1982).

Most of the above formulations are very general, since the specification of an underlying return-generating process is not required. However, we are concerned about their possible lack of power, made worse in this context by the fact that covariance and coskewness with the market are almost perfectly collinear across portfolios. Of course in the extreme case, in which market covariance is proportional to market coskewness, it will be impossible to identify covariance and coskewness premia separately. Therefore, in order to identify and accurately measure the contribution of coskewness, in this paper we propose an approach (see also Barone-Adesi 1985) based on the prior specification of an appropriate return-generating process: the quadratic market model. 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 measure the marginal contribution of coskewness to expected excess returns. Since market returns and the square of the market returns are almost orthogonal regressors, we obtain a precise test of the significance of quadratic coefficients. In addition, this framework allows us to test an asset pricing model with coskewness by checking the restrictions it imposes on the coefficients of the quadratic market model. The specification of a return-generating process provides more powerful tests as confirmed in a series of Monte Carlo simulations (see Section 5).

In addition to evaluating asset pricing models that include coskewness, it is also important to investigate the consequences on asset pricing tests when coskewness is erroneously omitted. We consider the possibility that portfolio characteristics such as size are empirically found to explain expected excess returns because of the omission of a truly priced factor, namely coskewness. To explain this problem, let us assume that coskewness is truly priced, but it is omitted in an 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, because it proxies for omitted coskewness. In our empirical application (see Section 4) we actually find that coskewness and firm size are correlated. This finding suggests that the empirically observed relation between size and assets excess returns may be explained by the omission of a systematic risk factor, namely market coskewness (see also Harvey and Siddique 2000, p. 1281).

The article is organized as follows. Section 2 introduces the quadratic market model. An asset pricing model including coskewness is derived using arbitrage pricing, and the testing of various related statistical hypotheses is discussed. Section 3 reports estimators and test statistics used in the empirical part of the paper. Section 4 describes the data, and reports empirical results. Section 5 provides Monte Carlo simulations for investigating the finite sample properties of our test statistics, and Section 6 concludes.

2 ASSET PRICING MODELS WITH COSKEWNESS

In this Section we introduce the econometric specifications considered in this paper. In turn, we describe the return-generating process, we derive the corresponding restricted equilibrium models, and finally we compare our approach with a GMM framework.

2.1 The Quadratic Market Model

Factor models are amongst the most widely used return-generating processes in financial econometrics. They explain co-movements in asset returns as arising from the common effect of a (small) number of underlying variables, called factors (see for instance Campbell, Lo, and MacKinlay 1987; Gourieroux, Jasiak 2001). In this paper, a linear two-factor model, the quadratic market model, is used as a return-generating process. Market returns and the square of the market returns are the 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, portfolio and market excess returns are defined by: $r_t = R_t - R_{F,t}\iota$, $r_{M,t} = R_{M,t} - R_{F,t}$, respectively, where ι is a $N \times 1$ vector of ones. Similarly, the excess squared market return is defined by: $q_{M,t} = R^2_{M,t} - R_{F,t}$. The quadratic market model is specified by:

$$\begin{array}{rl} & r_t = \alpha + \beta r_{M,t} + \gamma q_{M,t} + \varepsilon_t,\ t = 1, ..., T, \\ \mathcal{H}_F: & \gamma \in \mathbb{R}^N \end{array} \tag{1}$$

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:

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

with $\underline{R_{M,t}}$ and $\underline{R_{F,t}}$ denoting all present and past values of $R_{M,t}$ and $R_{F,t}$.

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

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

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. More specifically, 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 therefore show 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 finding is supported by our empirical investigations in Section 4, where, in accordance with the results of Harvey and Siddique (2000), we find that portfolios formed by assets of small firms tend to have 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 important risk characteristic: an asset which has positive coskewness with the market diminishes the sensitivity of a portfolio to large absolute market returns. Therefore, 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:

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

where ϵ_t (respectively $\epsilon_{q,t}$) are the residuals from a theoretical regression of portfolio returns R_t (market square returns $R^2_{M,t}$, respectively) on a constant and market return $R_{M,t}$. Since coefficients γ are proportional to $cov\left[\epsilon_t, R^2_{M,t}\right]$, we can use the estimate of γ in model (1) to investigate the coskewness properties of the N portfolios in the sample. Moreover, although γ does not correspond exactly to the usual probabilistic definition of market coskewness, coefficient γ is a very good proxy for $cov\left(r_t, R^2_{M,t}\right)/V\left(R^2_{M,t}\right)$, as pointed out in Kraus and Litzenberger (1976). Within our sample the approximation error is smaller than 1% (see Appendix A). Finally, the statistical (joint) significance of coskewness coefficient γ can be assessed by testing the null hypothesis $\mathcal{H}^*_F$ against the alternative $\mathcal{H}_F$.

2.2 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 for example by arbitrage pricing (Ross 1976; Chamberlain and Rothschild 1983). The arbitrage pricing theory (APT) implies that expected excess returns of assets following the factor model (1) satisfy the restriction (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 factors $r_{M,t}$ 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 with the 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 $q_{M,t}$ 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 equation (1) and substituting equations (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:

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

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

When model (4) is not supported by data, there exists an additional component $\widetilde{\alpha}$ (a $N \times 1$ vector) in expected excess returns, which cannot be fully related to market risk and coskewness risk: $E(r_t) = \beta\lambda_1 + \gamma\lambda_2 + \widetilde{\alpha}$. In this case, intercepts α of model (1) satisfy: $\alpha = \vartheta\gamma + \widetilde{\alpha}$. 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 our portfolios to this factor. Furthermore, variables representing portfolio characteristics, which are correlated with $\widetilde{\alpha}$ across portfolios, will have spurious explanatory power for expected excess returns, since they are a proxy for the sensitivities to the omitted factor. A case of particular interest is when $\widetilde{\alpha}$ is homogeneous across assets: $\widetilde{\alpha} = \lambda_0\iota$, where λ_0 is a scalar, that is:

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

corresponding to the following specification:

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

Specification (8) corresponds to the case where the factor omitted in model (4) has homogeneous sensitivities across portfolios. From equation (7), λ_0 may be interpreted as the expected excess returns of a portfolio 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 (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 faced by investors. With reference to the observed empirical regularities and model misspecifications mentioned in the Introduction, 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 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 $\mathcal{H}_2$.

2.3 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, the asset pricing model with coskewness (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 non-linear in the market returns. Thus, in the GMM approach, the derivation and testing of the orthogonality condition (9) do not require a prior specification of a data generating process.

More recently, in a conditional GMM framework, Dittmar (2002) uses a stochastic discount factor model embodying both quadratic and cubic terms. 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, the main contribution of our paper, beyond the results obtained by Harvey and Siddique (2000) and Dittmar

(2002), 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. a GMM test).

3 ESTIMATORS AND TEST STATISTICS

In this Section we derive the estimators and test statistics used in our empirical applications. Following an approach widely adopted in the literature (see for instance Campbell, Lo, and MacKinlay 1997; Gourieroux and Jasiak 2001), we consider the general framework of Pseudo Maximum Likelihood (PML) methods. We derive the statistical properties of the estimators and test statistics within the different coskewness asset pricing models presented in Sections 2.1 and 2.2. For completeness, full derivations are provided in the Appendices.

3.1 The Pseudo Maximum Likelihood Estimator

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

$$\begin{aligned} E\left[\varepsilon_t | \underline{\varepsilon_{t-1}}, \underline{R_{M,t}}, \underline{R_{F,t}}\right] &= 0, \\ E\left[\varepsilon_t \varepsilon_t^{\prime} | \underline{\varepsilon_{t-1}}, \underline{R_{M,t}}, \underline{R_{F,t}}\right] &= \Sigma, \end{aligned} \tag{10}$$

where Σ is a positive definite $N \times N$ matrix. The factor $f_t = (r_{M,t}, q_{M,t})^{\prime}$ is supposed to be exogenous in the sense of Engle, Hendry and Richard (1988). The expectation and the variance-covariance matrix of factor f_t are denoted by μ and Σ_f, respectively. Statistical inference in the asset pricing models presented in Section 2 is conveniently cast in the general framework of 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 criterion $L_T(\theta)$ is a (conditional) pseudo-loglikelihood. More specifically, $L_T(\theta)$ is the (conditional) loglikelihood of the model when we adopt a given conditional distribution for error ε_t, which satisfies (10) and is such that the resulting pseudo true density of the model is exponential quadratic. Under standard regularity assumptions, the PML estimator $\widehat{\theta}$ is consistent for any chosen conditional distribution of error ε_t satisfying the above conditions (see references above). Estimator $\widehat{\theta}$ 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. Finally, since the PML estimator is based on the maximization of a statistical criterion, hypothesis testing can be conducted by the usual general asymptotic tests.

In what follows, we will systematically analyze in the PML framework the alternative specifications introduced in Section 2.

3.2 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. Let θ denote the parameters of interest in model (1), where $vech(\Sigma)$ is a $\frac{(N+1)N}{2} \times 1$ vector representation of Σ containing only elements on and above the main diagonal,

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

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 model (1). Let B denote the $N \times 3$ matrix defined by $B = [\alpha \ \beta \ \gamma]$. The PML estimator $\widehat{B} = \left[\widehat{\alpha} \ \widehat{\beta} \ \widehat{\gamma}\right]$ is consistent when $T \to \infty$ and its asymptotic distribution is given by:

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

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

Let us now consider the test of the (joint) statistical significance of the coskewness coefficients γ, that is the test of hypothesis $\mathcal{H}_F^*$: $\gamma = 0$, against $\mathcal{H}_F$. This test can be easily performed computing a Wald statistics, which is given by:

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

where upper indices in a matrix denote elements of the inverse. Statistics ξ_T^{F*} is asymptotically $\chi^2(p)$-distributed, with $p = N$, when $T \longrightarrow \infty$.

3.3 Restricted Equilibrium Models

Let us now consider the constrained models (6) and (8) derived by arbitrage equilibrium. The estimation of these models is less simple 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 $\widehat{\theta}$ is given by the following system of implicit equations (see Appendix B):

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

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

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

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 given by:

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

Estimator $\left(\widehat{\beta}', \widehat{\gamma}'\right)'$ 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 equation (16); c) computing $\left(\widehat{\vartheta}, \widehat{\lambda}_0\right)'$ from equation (17) using new estimates for $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{Z}$; d) computing $\widehat{\Sigma}$ from equation (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 equation (1) are used).

The asymptotic distributions of the PML estimator are reported in Appendix B. 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 it satisfies the conditions for PML estimation. The results for the constrained PML estimation of model (6) follow by setting $\lambda_0 = 0$, $\widehat{Z} = \widehat{\gamma}$, and deleting the vector ι.

Let us now consider the problem of testing hypotheses $\mathcal{H}_1$ and $\mathcal{H}_2$, corresponding to models (6) and (8) respectively, against the alternative $\mathcal{H}_F$. If θ denotes the parameter of model (1), each of these two hypotheses can be written in mixed form:

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

for an appropriate vector function g with values in $\mathbb{R}^r$ and suitable dimensions q and r. Let us assume 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. The test of 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). More specifically the test is based on the following statistics:

$$\xi_T = \arg\min_a Tg(\widehat{\theta}, a)'\widehat{S}g(\widehat{\theta}, a),$$

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 $\chi^2(r-p)$-distributed and is asymptotically equivalent to the other asymptotic tests.

By applying these general results, we deduce the ALS statistics for testing the hypotheses $\mathcal{H}_1$ and $\mathcal{H}_2$ against the alternative $\mathcal{H}_F$ (see Appendix C for a full derivation). The hypothesis $\mathcal{H}_1$ against $\mathcal{H}_F$ is tested by the statistics:

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

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 ALS statistics for testing hypothesis $\mathcal{H}_2$ against $\mathcal{H}_F$ is given by:

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

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

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

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

4 EMPIRICAL RESULTS

In this Section we report the results of our empirical application, performed on monthly returns of stock portfolios. We first estimate the quadratic market model, and then test the different associated asset pricing models with coskewness. Finally we investigate the consequences of erroneously neglecting coskewness when testing asset pricing models. The Section begins with a brief description of the data.

4.1 Data Description

Our dataset consists of 450 (percentage) monthly returns of the 10 stock portfolios formed according to size by French, for the period July 1963-December 2000. Data are available from the web site:
http://web.mit.edu/kfrench/www/data_library.html, in the file "Portfolios Formed on Size". The portfolios are constructed at the end of June each year, using June market equity data and NYSE breakpoints. The portfolios from 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 being 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. The market return and risk-free return are available from the web 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.

4.2 Results

4.2.1 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 1 and 2, respectively.

[**Insert somewhere here Tables 1 and 2**]

As explained in Section 3.2, these estimates are obtained by OLS regressions, performed equation by equation on system (1). As expected, the beta coefficients are strongly significant for all portfolios, with smaller portfolios having higher betas in general. From the estimates of the γ parameter, we see that small portfolios have significantly negative coefficients of market coskewness (for instance $\gamma = -0.017$ for the smallest portfolio). Coskewness coefficients are significantly positive for the two largest portfolios ($\gamma = 0.003$ for the largest portfolio). In particular, we observe that the β and γ coefficients are strongly correlated across portfolios. We can test for joint significance of the coskewness parameter γ by using the Wald statistics ξ_T^{F*} in (14). The statistics ξ_T^{F*} assumes the value: $\xi_T^{F*} = 35.34$, which is strongly significant at the 5 percent level, because its critical value is $\chi^2_{0.05}(10) = 18.31$. Finally, from Table 2, we also see that smaller portfolios are characterized by larger variances of the residual error terms.

We performed several specification tests of the functional form of the mean portfolios return in equation (1). First, we estimated a factor SUR model including also a cubic power of market returns, $R_{M,t}^3 - R_{F,t}$, as a factor in addition to the constant, market excess returns and market squared excess returns. The cubic factor was found not to be significant for all portfolios. Furthermore, in order to test for more general forms of misspecifications in the mean, we performed Ramsey (1969) Reset Test on each portfolio, 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 was accepted for all portfolios in our tests.

From the point of view of our analysis, one central result from Table 1 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 assumed in the market model), generating 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 firm portfolios, having negative market coskewness, are exposed to a source of risk additional to market risk, related to the occurrence of large absolute market returns. In addition, as we have already seen, market model (2), when tested against quadratic market model (1), is rejected with a largely significant Wald statistics. In the light of our findings, we conclude that the extension of the return-generating process to include the squared market return is valuable.

4.2.2 Restricted Equilibrium Models

Let us now investigate market coskewness in the context of models which are consistent with arbitrage pricing. To this end we consider constrained PML estimation of models (6) and (8). Specification (6) is obtained from the quadratic market model after imposing restrictions from the asset pricing model (4). Specification (8) instead allows for a homogeneous additional constant in expected excess returns. The corresponding PML estimators are obtained from the algorithm based on equations (16) to (18), as reported in Section 3.3. The results for model (6) are reported in Table 3 and those for model (8) in Table 4.

[**Insert somewhere here Tables 3 and 4**]

The point estimates and standard errors of parameters β and γ are similar in the two models. Their values are close to those obtained from quadratic market model (1). In particular, the estimates of parameter γ confirm that small (large) portfolios have significantly negative (positive) coskewness coefficients. 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 either model. However, the estimate in model (8), $\widehat{\lambda}_2 = -7.439$, has at least the expected negative sign. Using this estimate, we deduce that, for a portfolio with coskewness $\gamma = -0.01$ (a moderately-sized portfolio, such as portfolio 3 or 4), the coskewness contribution to the expected excess return on an annual percentage basis is approximately 0.9. This value increases to 1.5 for the smallest portfolio in our data set.

We test the empirical validity of 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 some evidence that asset pricing model (4) may not be satisfied in our sample. In other words, an additional component, other than covariance and coskewness to market, may be present in expected excess returns. In order to test for the homogeneity of this component across assets, we test hypothesis $\mathcal{H}_2$ against $\mathcal{H}_F$. The test statistics ξ_T^2 in equation (22) assumes the value $\xi_T^2 = 5.32$, well below the critical value $\chi^2_{0.05}(8) = 15.51$. A more powerful test of asset pricing model (4) should be provided by testing hypothesis $\mathcal{H}_1$ against the alternative, $\mathcal{H}_2$. This test is performed by the simple t-test of significance of λ_0. From Table 4 we see that $\mathcal{H}_1$ is quite clearly rejected. This confirms that asset pricing model (4) may not be supported by our data. However, since $\mathcal{H}_2$ is not rejected, it follows that, if the additional component which is unexplained by model (4) comes from an omitted factor, its sensitivities should be homogeneous across portfolios in our sample. We conclude that size is unlikely to 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 half the contribution due to coskewness for portfolios of modest size. As explained in Section 2.2, $\lambda_0 > 0$ may be due to the use of a risk-free rate lower than the actual rate faced by investors.

4.2.3 Misspecification from Neglected Coskewness

As already mentioned in Section 2, we are also interested in investigating the consequences on asset pricing tests of erroneously neglecting coskewness. The results presented so far suggest that market model (2) is misspecified, since it does not take into account the quadratic market return. Indeed, when tested against quadratic market model (1), it is strongly rejected. For comparison, we report the estimates of parameters α and β in market model (2) in Table 5.

[Insert somewhere here Table 5]

The β coefficients in Table 5 are close to those obtained in the quadratic market model in Table 1. Therefore, neglecting the quadratic market returns does not seem to have dramatic consequences for the estimation of parameter β. However, we expect 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 feature suggests that size can have a spurious explanatory power in the cross-section of expected excess returns since it is a proxy for omitted coskewness. Therefore, as anticipated in Section 2, the ability of size to explain expected excess returns could be due to misspecification of models neglecting coskewness risk.

Finally, it is interesting to compare our findings 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 over time, with assets having negative coskewness commanding, not surprisingly, higher expected returns. 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, according to Kan and Zhang (1999a,b)'s criterion the size effect is more likely explained by neglected coskewness.

5 MONTE CARLO SIMULATIONS

In this final Section we report the results of a series of Monte Carlo simulations undertaken to assess the importance of specifying the return-generating process to obtain reliably powerful statistical tests. 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 equation (21), which tests model (4) by the restrictions imposed on the return-generating process (1), and ii) a GMM test statistics ξ_T^{GMM}, which tests model (4) through the orthogonality conditions (9). In addition, we investigate the effects on the ALS statistics ξ_T^1 induced by the non-normality of errors ε_t or by the misspecification of the return-generating process (1).

5.1 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:

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

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 1, matrix Σ is taken from Table 2, $\vartheta = -14.995$ from Table 3, $\mu_m = 0.52$, $\sigma_m = 4.41$, 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 $\lambda_0 = 0$, quadratic equilibrium model (4) is satisfied. When $\lambda_0 \neq 0$, equilibrium model (4) is not correctly specified, and the misspecification is in the form of an additional component, which is homogeneous across portfolios, corresponding to model (8). However, for any value of λ_0, quadratic market model (1) is well specified.

We perform a Monte Carlo simulation $(10,000$ 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 6.

[Insert somewhere here Table 6]

The second row, $\lambda_0 = 0$, reports the empirical test sizes. Both statistics control size quite well in finite samples, 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 empirical analysis in Section 4. Table 6 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 given by (1), whereas the alternative for the GMM statistics ξ_T^{GMM} is left unspecified.

5.2 Experiment 2

Under DGP1, residuals ε_t are normal. When residuals ε_t are not normal, the alternative used by the ALS statistics ξ_T^1 [i.e. 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 of non-normality of residuals ε_t on the ALS test statistics. The data-generating process used in this experiment, called DGP2, is equal to DGP1 but residuals ε_t follow a multivariate t-distribution with $df = 5$ degrees of freedom. The correlation matrix is chosen so that the resulting variance of residuals ε_t is the same as under DGP1. The rejection frequencies of this Monte Carlo simulation ($10,000$ replications) for the ALS statistics ξ_T^1 are reported in Table 7.

[**Insert somewhere here Table 7**]

The ALS statistics appears to be only slightly oversized. As expected, 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 non-normality of the residuals.

5.3 Experiment 3

In the Monte Carlo 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 hypothesis in the form of conditional heteroscedasticity of errors ε_t. We thus consider two data-generating processes having the same unconditional variance of the residuals ε_t, but so that the residuals ε_t are conditionally heteroscedastic in one case, and homoscedatic in the other. Specifically, DGP3 is the same as DGP1, but innovations ε_t follow a conditionally normal, multivariate ARCH(1) process without cross effects:

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

Matrix $= [\omega_{ij}]$ is chosen as in Table 2, and $\rho = 0.2$. DGP4 is similar to DGP1, with i.i.d. normal innovations with the same unconditional variance matrix as innovations 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 8.

[**Insert somewhere here Table 8**]

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

6 CONCLUSIONS

In this paper we have considered market coskewness and investigated its implications for testing asset pricing models. By estimating a quadratic market model, which includes the market returns and the square of the market returns as the two factors, we showed that portfolios of small (large) firms have negative (positive) coskewness with market returns. This finding implies that small firm portfolios are exposed to a source of risk, that is market coskewness, which is different from the usual market beta and arises from negative covariance with large absolute market returns. The coskewness coefficients of the portfolios in our sample are shown to be statistically significant. This finding rejects the usual market model and demonstrates the validity of the quadratic market model as a possible extension.

The analysis of the premium in expected excess returns induced by market coskewness requires the specification of appropriate asset pricing models including coskewness among the rewarded factors. In order to obtain testing methodologies that are more powerful compared to a GMM approach, we tested asset pricing models including coskewness through the restrictions they impose on the quadratic market model. We used asymptotic test statistics whose finite sample properties are validated by means of a series of Monte Carlo simulations. We found evidence of a component in expected excess returns that is not explained by either covariance or coskewness with the market. However, this unexplained component is relatively small and is consistent for instance with a minor misspecification of the risk-free rate.

More importantly, the unexplained component is homogeneous across portfolios. This finding implies that additional variables, representing portfolios characteristics such as firm size, have no explanatory power for expected excess returns once coskewness is taken into account.

Finally, the homogeneity property of the unexplained component in expected excess returns is not satisfied when coskewness is neglected. Therefore our results have important implications for testing methodologies, showing that neglecting coskewness risk can cause misleading inference. Indeed, since coskewness is positively correlated with size, a possible justification for the anomalous explanatory power of size in the cross-section of expected returns is that it is a proxy 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 in a very long time.

ACKNOWLEDGMENTS

We wish to thank the participants in the 9th Conference on Panel Data, the 10th Annual Conference of the European Financial Management Association, the Conference on "Multimoment Capital Asset Pricing Models and Related Topics", the INQUIRE UK 14th Annual Seminar on "Beyond Mean-Variance: Do Higher Moments Matter ?", the 2004 ASSA Meeting and P. Balestra, S. Cain-Polli, J. Chen, C. R. Harvey, R. Jagannathan, R. Morck, M. Rockinger, T. Wansbeek, C. Zhang, two anonymous referees and the Editor, Eric Ghysels. All of them have contributed, through discussion, very helpful comments and suggestions, to improving this paper. The usual disclaimer applies. Thanks are also due to 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.

APPENDIX A: RELATIONSHIP BETWENN CROSS MOMENT COSKEWNESS AND THE γ_i PARAMETER

In this Appendix we show how the parameter γ_i relates to the coskewness term of our quadratic market model. We also report error estimates when the cross moment coskewness is approximated by the parameter γ_i.

Our basic model is [see (1)]:

$$\begin{aligned} r_{t,i} &= \alpha_i + \beta_i(R_{M,t} - R_{F,t}) + \gamma_i(R_{M,t}^2 - R_{F,t}) + \varepsilon_{i,t}, \\ \text{where } r_{t,i} &= R_{t,i} - R_{F,t}. \end{aligned}$$

The probabilistic measure of coskewness is defined by:

$$cov(r_{t,i}, R_{M,t}^2) = E[r_{t,i}R_{M,t}^2] - E[r_{t,i}]E[R_{M,t}^2],$$

which can be rewritten as:

$$cov(r_{t,i}, R_{M,t}^2) = \beta_i(E[R_{M,t}^3] - E[R_{M,t}^2]E[R_{M,t}]) + \gamma_i Var[R_{M,t}^2] + E[R_{M,t}^2]E[\varepsilon_{t,i}].$$

In the final equation, the first term is a measure of the market asymmetry, the second is essentially our measure of coskewness and the final term is equal to zero by assumption (10). Evidently, our approach considers the second term only. However, our claim is motivated by the negligible effects of the first term. In fact, for values $\beta = 1$ and $\gamma_i = -0.01$, which are representative for small firm portfolios, the first term is 0.1, while the second is -15. If γ_i is equal to 0.003, as in large firm portfolios, then the terms are equal to 0.1 and 4 respectively. Finally, if the portfolio has a $\gamma_i = 0$, the second term is also 0.

We are greatly indebted to one of the referees, whose comments highlighted this point.

APPENDIX B: 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 text) 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 pseudo-loglikelihood are given by:

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

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

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

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 $\left(\beta',\gamma'\right)'$, $\left(\vartheta,\lambda_0\right)'$]:

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

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

$$
V_{as}\left[\sqrt{T}\left(\widehat{\theta}-\theta_0\right)\right] = J_0^{-1}I_0J_0^{-1} = \begin{bmatrix} J_0^{*-1} & J_0^{*-1}\eta S \\ S'\eta' J_0^{*-1} & K \end{bmatrix}.
$$

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 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:

$$J_0^{*-1} = \left[\begin{array}{cc} J_0^{*11} & J_0^{*12} \\ J_0^{*21} & J_0^{*22} \end{array}\right],$$

where:

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

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

$$J_0^{*21} = J_0^{*12\prime}$$

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

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

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

where $L_T\left(\theta\right)$ 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 C: 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 ${}_0 = V_{as}\left(\sqrt{T}\left(\widehat{\theta} - \theta\right)\right)$. From (13) we get:

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

The test statistics follow.

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; Velu and Zhou 1999). These tests are asymptotically equivalent to the Asymptotic Least Squares tests and are used in our paper for their computational simplicity. An evaluation of the finite sample properties of the ALS test statistics is presented in Section 5.

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 Talwar, P. (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 MacKinlay, A.C. (1987). *The Econometrics of Financial Markets.* Princeton, New Jersey: Princeton University Press.

Chamberlain, G., and Rothschild, M.(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 the Cross Section of Equity Returns," *Journal of Finance,* 57, 368-403.

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

Fama, E., and French, K. R. (1995), "Size and Book-to-Market Factors in Earnings and Returns," *Journal of Finance,* 50, 131-55.

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

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

Gourieroux, C., Monfort, A., and Trognon, A. (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 Jagannathan, R. (1997), "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance,* 52, 557-590.

Harvey, C.R., and Siddique, A. (2000), "Conditional Skewness in Asset Pricing Tests," *Journal of Finance,* 55, 1263-1295.

Jagannathan, R., and Wang, Z. (1996), "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance,* 51, 3-53.

Jagannathan, R., and Wang, Z. (2001), "Empirical Evaluation of Asset Pricing Models: A Comparison of the SDF and Beta Methods," Working Paper, Northwestern University, Illinois.

Kan, R., and Zhang, C. (1999a), "GMM Test of Stochastic Discount Factor Models with Useless Factors," *Journal of Financial Economics,* 54, 103-27.

Kan, R., and Zhang, C. (1999b), "Two-Pass Tests of Asset Pricing Models with Useless Factors," *Journal of Finance,* 54, 203-35.

Kan, R., and Zhou, G. (1999), "A Critique of the Stochastic Discount Factor Methodology," *Journal of Finance,* 54, 1221-1248.

Kraus, A., and Litzenberger, R. (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.

Merton, R. C.(1973), "An Intertemporal Capital Asset Pricing Model," *Econometrica,* 41, 867-87.

Newey, W. K., and West, K.D. (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.

Rubinstein, M. E. (1973), "A Comparative Statics Analysis of Risk Premium," *The Journal of Business*, 46, 605-615.

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 Zhou, G. (1999), "Testing Multi-Beta Asset Pricing Models," *Journal of Empirical Finance*, 6, 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, 71-93.

Table 1. 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]

NOTE: This table 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:

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

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

Table 2. 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

NOTE: This table 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 $r_t = R_t - R_{F,t}\iota$, $r_{M,t} = R_{M,t} - R_{F,t}$, $q_{M,t} = R^2_{M,t} - R_{F,t}$. 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 3. 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)

NOTE: This table 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 4. 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)

NOTE: This table 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 5. 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)

NOTE: This table 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 6. 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

NOTE: This table 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

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

and

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

Parameters β and γ are the third and fourth columns respectively in Table 1, the matrix Σ is taken from Table 2, $\vartheta = -14.995$ from Table 3, $\mu_m = 0.52$, $\sigma_m = 4.41$, 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 7. 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

NOTE: This table 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 4), but the residuals ε_t follow a multivariate t-distribution with $df = 5$ degrees of freedom, and a correlation matrix such that the variance of ε_t is the same as under DGP1.

Table 8. 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

NOTE: This table reports the rejection frequencies of the ALS statistics ξ_T^1 [in (21)] for testing (4):

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

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 homoscedastic in the other. Specifically, DGP3 is the same as DGP1 (see Table 6), 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 2, and $\rho = 0.2$. DGP4 is similar to DGP1 (see Table 6), 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.