Option-Based Estimation of Co-Skewness and Co-Kurtosis Risk Premia

Peter Christoffersen
University of Toronto, CBS, and CREATES

Kris Jacobs
University of Houston and Tilburg University

Mehdi Karoui
McGill University and OMERS

February 19, 2014

Abstract

We show that the price of risk for equity factors that are nonlinear in the market return are readily obtained using index option prices. We apply this insight to the price of co-skewness and co-kurtosis risk. The price of co-skewness risk corresponds to the spread between the physical and the risk-neutral second moments, and the price of co-kurtosis risk corresponds to the spread between the physical and the risk-neutral third moments. Our option-based estimates of the prices of risk lead to reasonable values of the associated risk premia. An out-of-sample analysis of factor models with co-skewness and co-kurtosis risk indicates that the new estimates of the price of risk improve the models' performance. The models using higher-order market moments also robustly outperform standard competitors such as the CAPM and the Fama-French model.

JEL Classification: G12, G13, G17

Keywords: Co-skewness; co-kurtosis; risk premia; options; cross-section; out-of-sample.

1 Introduction

The specification and performance of factor models are of paramount importance for financial research and practice, and have been the subject of intense debate for a long time. The Capital Asset Pricing Model (CAPM) has been criticized from different angles, and although its performance improves substantially when evaluating the model conditionally rather than unconditionally, there is widespread consensus that models with better explanatory power are badly needed.

Many alternative models have been proposed over the past four decades, with limited success. One class of models attempts to find new factors using economic intuition or more formal economic modeling. The performance of these models in cross-sectional pricing has been rather disappointing. For instance, aggregate consumption, which is a state variable suggested by theory, has been shown to have limited explanatory power for the cross-section of stock returns. Another class of models constructs factors using a more reduced-form approach, partly based on well-documented stylized facts. The standard examples in this literature are the three-factor model of Fama and French (1993), which includes market, book-to-market and size factors, and the four-factor model suggested by Carhart (1997), which additionally includes a momentum factor. The cross-sectional explanatory power of these models is often judged as satisfactory, but the lack of economic and theoretical foundations is cause for concern.[1]

In view of the state of the literature, further evidence on the pricing of the cross-section of stock returns is therefore a priority. This paper contributes to a literature that goes back to Kraus and Litzenberger (1976), who argue that if investors care about portfolio skewness, co-skewness enters as a second pricing factor in addition to the market portfolio. This argument has later been applied to investor preferences over portfolio kurtosis, leading to co-kurtosis as an additional factor (see, for instance, Ang, Chen, and Xing (2006), Dittmar (2002), Guidolin and Timmermann (2008), and Scott and Horvath (1980)).[2] Despite several important contributions by among others Bansal and Viswanathan (1993), Leland (1997), Lim (1989), Harvey and Siddique (2000), and Dittmar (2002), and despite the theory's obvious intuitive appeal, there seems to be no widespread consensus on the importance of this literature for cross-sectional asset pricing.

[1] An extensive literature has sprung up that attempts to provide economic underpinnings for the Fama-French and Carhart factors. See for example Liew and Vassalou (2000) for a risk-based explanation, and Chan, Karceski, and Lakonishok (2003) for a behavioral explanation.

[2] See also Arditti (1967), Rubinstein (1976), and Golec and Tamarkin (1998) for related work.

One possible drawback of co-skewness and co-kurtosis as cross-sectional pricing factors is measurement. Most existing papers estimate and test the importance of co-skewness and co-kurtosis using two-stage cross-sectional regressions. For classical examples of such analyses of co-skewness risk, see for instance Kraus and Litzenberger 1976) and Harvey and Siddique (2000). This approach necessitates the estimation of co-skewness betas in a first stage. These betas are subsequently used in the second-stage cross-sectional regression. It is well-known that there may be biases in the estimation of betas in the first-stage regression, and these errors carry over in the second-stage cross-sectional regression. While these problems apply to virtually all implementations of cross-sectional models, including the CAPM, they may be especially serious in the case of co-skewness and co-kurtosis. The simple basic intuition is that the higher the moment, the more difficult it is to estimate precisely. This argument applies a fortiori to the estimation of co-measures of higher moments, such as co-skewness and co-kurtosis, and the betas for these factors. Therefore, errors in estimated betas may be large for these models, leading to biases in the cross-sectional estimation of the price of risk that are potentially much larger than in the competing case of the CAPM or the Fama-French three-factor model.

We propose a new strategy to estimating the price of co-skewness and co-kurtosis risk, which avoids the problems inherent in the second-stage cross-sectional regression. Our approach can also be used to estimate the price of other risks, provided that they are nonlinear functions of the market return. We derive our result based on the well-known representation of cross-sectional asset pricing models that relies on the stochastic discount factor or SDF (see Cochrane (2005)). The CAPM corresponds to the assumption of linearity of the SDF with respect to the market return. A quadratic SDF implies that investors require compensation not only for the exposure to market returns but also for the exposure to squared market returns, which leads to co-skewness risk aversion.[3] SDFs that are higher-order functions of the market return lead to progressively more complex co-movements with market returns as pricing factors.

The key difference between our approach and existing studies is that we explicitly impose restrictions on the pricing of both stocks and contingent claims. This allows us to derive explicit formulas for the time-varying price of risk for the exposure to any nonlinear function of the market return . For instance, for the case of co-skewness risk we show that the price of

[3] See Dittmar (2002) for an investigation of higher moments in cross-sectional pricing using this approach. See Bakshi, Madan, and Panayotov (2010) for evidence that pricing kernels are U-shaped as a function of market returns.

co-skewness risk corresponds to the spread between the physical and the risk-neutral second moment. Similarly, the price of co-kurtosis risk is given by the spread between the physical and the risk-neutral third moment. To provide intuition for this result, consider the special case where the SDF is a linear function of the market return, which corresponds to the CAPM. In this case, our general result shows that the price of risk can be estimated as the difference between the spread between the physical and risk-neutral first moment. This equals the market return minus the risk-free rate, which is of course the classical CAPM result.

We empirically investigate the performance of our approach for the pricing of co-skewness and co-kurtosis risk. Using monthly data for the period 1996-2012, we find that the price of co-skewness risk has the expected negative sign in every month in our sample, and the price of co-kurtosis risk has the expected positive sign in most months. On average, both estimated prices of risk are somewhat larger in absolute value than the traditional estimates obtained using a two-stage Fama-MacBeth approach. More importantly, while the average prices of risk obtained using the Fama-MacBeth approach have the theoretically anticipated signs on average, they are often estimated with the opposite sign. We evaluate the cross-sectional performance of our newly proposed estimates out-of-sample, and find that they outperform implementations of the CAPM and the Fama-Frech three factor model that use cross-sectional regressions to estimate the price of risk.

The paper proceeds as follows. Section 2 describes our alternative approach to the measurement of (nonlinear) market risk. Section 3 estimates the physical and risk-neutral higher moments needed to compute the price of risk in our approach. Section 4 presents cross-sectional stock return evidence, including an extensive out-of-sample analysis. Section 5 contains robustness analyses, and Section 6 concludes.

2 Measuring Market Risks: An Option-Based Approach

In this section we provide an overview of multifactor linear models in which cross-sectional differences in expected returns between assets are determined by their exposure to risk factors that are nonlinear functions of the market return. This setting corresponds to assuming SDFs that are nonlinear in the market return. We proceed to propose an option-based approach to measuring the price of risk for these types of exposures. We investigate two special cases that are of significant empirical interest: exposure to the squared market return R_m^2, which captures co-skewness risk; and exposure to the third power of the market return R_m^3, which captures co-kurtosis risk.

2.1 Measuring Co-Skewness Risk

Before we introduce the general case, we first discuss two specific examples to provide more intuition for our approach. We begin with co-skewness risk. Let m_{t+1} denote the stochastic discount factor

$$m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R^2_{m,t+1} - E_t^P(R^2_{m,t+1})\right), \tag{1}$$

where $R_{m,}$ denotes the stock market return, and $E_t^P(.)$ denotes the expectation under the physical probability measure. Similar to Harvey and Siddique (2000, henceforth HS), our setup is based on the assumption of a quadratic SDF. As explained by HS (2000), a quadratic SDF is consistent with several utility-based asset pricing models. The performance of quadratic pricing kernels is studied in Bansal and Viswanathan (1993) and Chabi-Yo (2008).

Given this SDF, we can establish pricing restrictions on any asset return. The key feature of our approach is that we jointly consider theoretical restrictions on stocks and contingent claims, whereas the existing cross-sectional asset pricing literature focuses exclusively on the underlying assets. Our approach enables the specification of new estimators for the price of co-skewness risk which can be easily implemented using short data windows.

Let us denote by R_j the return on a stock j and by R_i the return on a contingent claim on the stock i. The existing literature contains several measures of co-skewness risk, which all capture covariation between the stock return and the squared market return. Kraus and Litzenberger (1976, henceforth KL) define co-skewness risk by $\frac{E^P\left[(R_j-\overline{R}_j)(R_m-\overline{R}_m)^2\right]}{E^P\left[(R_m-\overline{R}_m)^3\right]}$.

HS (2000) mainly focus on $cov(R_j, R_m^2)$ in their theoretical model. In the empirical analysis, they consider four different co-skewness measures, and we consider one of these in our empirical work below. We will report on the measure in HS (2000) that is similar to the one in KL (1976), namely

$$\beta_j^{SKD} = \frac{E^P\left[(\epsilon_j \epsilon_m^2\right]}{\sqrt{E^P\left[\epsilon_j^2\right] E^P\left[\epsilon_m^2\right]}}, \tag{2}$$

where ϵ_j is the residual from the regression of the asset's return on the market return.

Our measure of co-skewness risk β_j^{COSK} (β_i^{COSK}) of stock j (contingent claim i) is its beta with respect to R_m^2 in a multivariate regression. This measure allows for mathematical tractability in the derivation of the price of risk as shown in the following proposition. The proposition presents the pricing implications of the SDF defined in equation (1).

Proposition 1 *If the stochastic discount factor (SDF) has the following form:*

$$m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right),$$

then the cross-sectional pricing restrictions are

$$E_t^P(R_{j,t+1}) - R_f = \lambda_t^{MKT}\beta_{j,t}^{MKT} + \lambda_t^{COSK}\beta_{j,t}^{COSK}, \tag{3}$$

and

$$E_t^P(R_{i,t+1}) - R_f = \lambda_t^{MKT}\beta_{i,t}^{MKT} + \lambda_t^{COSK}\beta_{i,t}^{COSK}, \tag{4}$$

where β_t^{MKT} *and* β_t^{COSK} *are from the projection of the asset returns on* $R_{m,t+1}$*, and* $R_{m,t+1}^2$*. The price of covariance risk,* λ_t^{MKT}*, is*

$$\lambda_t^{MKT} = E_t^P(R_{m,t+1}) - R_f, \tag{5}$$

and the price of co-skewness risk, λ_t^{COSK}*, is*

$$\lambda_t^{COSK} = E_t^P(R_{m,t+1}^2) - E_t^Q(R_{m,t+1}^2). \tag{6}$$

where $E_t^P(.)$ *and* $E_t^Q(.)$ *denote the expectation under the physical and risk-neutral probability measures, respectively.*

Proof. See Appendix A. ■

Proposition 1 shows that the price of co-skewness risk corresponds to the spread between the physical and the risk-neutral second moments for the market return. Unlike other moments, the second moment is fairly easy to estimate under both the physical and risk-neutral probability measures. The literature contains a wealth of robust approaches for modeling the physical volatility of stock returns. The risk-neutral moment can be estimated from option market data either by the implied volatility of option pricing models, or alternatively using a model-free approach based as in Bakshi and Madan (2000) and Bakshi, Kapadia and Madan (2003).

A number of existing studies relate the volatility spread to risk aversion (see Bakshi and Madan (2006)) or the price of correlation risk (see Driessen, Maenhout and Vilkov (2009)).

Proposition 1 shows that if the pricing kernel is quadratic, then the volatility spread is equal to the price of co-skewness risk.

Proposition 1 allows for separate identification of the price of covariance (λ_t^{MKT}) and co-skewness (λ_t^{COSK}) risk. Note that this result is simply an application of the general result that if the factor is a portfolio, then the expected return on the factor is equal to the factor risk premium. Importantly, the result holds regardless of assumptions on other risk factors. This is in stark contrast with risk premia estimated from two-pass cross-sectional regressions for which the empirical results depend on the other risk factors considered in the regression. Our approach also has the advantage of easily capturing time variation in risk premia.

While our approach to estimating the price of co-skewness risk is different from the existing literature and the betas are defined differently, the implications for the risk premia on the assets are of course the same. Using the fact that $E_t^P(R_{m,t+1}) - R_f = \lambda_t^{MKT}$ and $E_t^P(R_{m,t+1}^2) - E_t^Q(R_{m,t+1}^2) = \lambda_t^{COSK}$, we can re-write equation (3) of proposition 1 as follows

$$E_t^P(R_{j,t+1}) - R_f = \beta_{j,t}^{MKT}\left[E_t^P(R_{m,t+1}) - R_f\right] + \beta_{j,t}^{COSK}\left[E_t^P(R_{m,t+1}^2) - E_t^Q(R_{m,t+1}^2)\right], \quad (7)$$

which can also be written as

$$E_t^P(R_{j,t+1}) - R_f = c_t + \beta_{j,t}^{MKT} E_t^P\left(R_{m,t+1}\right) + \beta_{j,t}^{COSK} E_t^P\left(R_{m,t+1}^2\right), \quad (8)$$

where $c_t = -\beta_{j,t}^{MKT} R_f - \beta_{j,t}^{COSK} E_t^Q(R_{m,t+1}^2)$. Equation (8) shows the link between our method and the approaches in KL (1976) and HS (2000). It is equivalent to equation (6) of KL (1976) and equation (8) of HS (2000).

The crucial difference between our approach and the one in KL (1976) and HS (2000) is that we explicitly impose the pricing restrictions on contingent claims. This additional restriction leads to a very simple estimator of the price of risk.

The existing empirical evidence clearly indicates that risk-neutral variance is larger than physical variance, therefore suggesting a negative price of co-skewness risk. See for instance Bakshi and Madan (2006), Bollerslev, Tauchen, and Zhou (2009), Carr and Wu (2009), and Jackwerth and Rubinstein (1996). A negative price of risk is consistent with theory. Assets with lower (more negative) co-skewness decrease the total skewness of the portfolio and increase the likelihood of extreme losses. Assets with lower co-skewness are thus perceived by investors to be riskier and should command higher risk premiums.

2.2 Measuring Co-Kurtosis Risk

A natural extension of the quadratic pricing kernel discussed in the previous section is the cubic pricing kernel studied in Dittmar (2002), given by

$$\begin{aligned} m_{t+1} &= a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right) \\ &\quad + b_{5,t}\left(R_{m,t+1}^3 - E_t^P(R_{m,t+1}^3)\right). \end{aligned} \tag{9}$$

A cubic pricing kernel is consistent with investors' preferences for higher order moments, specifically skewness and kurtosis. See Dittmar (2002) and HS (2000) for more details. As before, we first make an assumption on the shape of the SDF and then derive pricing restrictions. In this case, the expected excess return on any asset will be related to co-kurtosis risk, in addition to covariance risk and co-skewness risk. As explained by Dittmar (2002), kurtosis measures the likelihood of extreme values and co-kurtosis captures the sensitivity of asset returns to extreme market return realizations. If investors are averse to extreme values, they require higher compensation for assets with higher co-kurtosis risk, meaning that the price of co-kurtosis risk should be positive. See Guidolin and Timmermann (2008) and Scott and Horvath (1980) for a more detailed discussion. Similar to co-skewness risk, co-kurtosis risk has been defined in various ways in previous studies. For instance, Ang, Chen and Xing (2006) measure co-kurtosis risk using $\frac{E^P\left[(R_j-\overline{R}_j)(R_m-\overline{R}_m)^3\right]}{\sqrt{E^P\left[(R_j-\overline{R}_j)^2\right]}\left(E^P\left[(R_m-\overline{R}_m)^2\right]\right)^{3/2}}$, and Guidolin and Timmermann (2008) use $cov(R_j, R_m^3)$. In this paper, we measure co-kurtosis risk by the return's beta with respect to the cubic market return R_m^3. We denote the co-kurtosis beta of a stock j (contingent claim i) by $\beta_{j,t}^{COKU}$ ($\beta_{i,t}^{COKU}$).

The following proposition presents the estimator for the co-kurtosis price of risk and the cross-sectional pricing restrictions.

Proposition 2 *If the stochastic discount factor (SDF) has the following form:*

$$m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right),$$

then the cross-sectional restrictions are

$$E_t^P(R_{j,t+1}) - R_f = \lambda_t^{MKT}\beta_{j,t}^{MKT} + \lambda_t^{COSK}\beta_{j,t}^{COSK} + \lambda_t^{COKU}\beta_{j,t}^{COKU}, \tag{10}$$

and

$$E_t^P(R_{i,t+1}) - R_f = \lambda_t^{MKT}\beta_{i,t}^{MKT} + \lambda_t^{COSK}\beta_{i,t}^{COSK} + \lambda_t^{COKU}\beta_{i,t}^{COKU}, \tag{11}$$

where β_t^{MKT}, β_t^{COSK}, *and* β_t^{COKU} *are from the projection of asset returns on* $R_{m,t+1}$, $R^2_{m,t+1}$ *and* $R^3_{m,t+1}$, *respectively. The price of covariance,* λ_t^{MKT}, *and co-skewness risk* λ_t^{COSK} *are*

$$\lambda_t^{MKT} = E_t^P(R_{m,t+1}) - R_f, \tag{12}$$

$$\lambda_t^{COSK} = E_t^P(R^2_{m,t+1}) - E_t^Q(R^2_{m,t+1}), \tag{13}$$

and the price of co-kurtosis risk, λ_t^{COKU}, *is*

$$\lambda_t^{COKU} = E_t^P(R^3_{m,t+1}) - E_t^Q(R^3_{m,t+1}), \tag{14}$$

where $E_t^P(.)$ *and* $E_t^Q(.)$ *denote the expectation under the physical respectively risk-neutral probability measure.*

Proof. See Appendix A. ■

Proposition 2 shows that the price of co-kurtosis risk is equal to the spread between the market physical and risk-neutral third moments. Clearly third moments are harder to estimate than second moments. Nevertheless, existing evidence (see for instance Bakshi, Kapadia and Madan (2003)) indicates that the risk-neutral distribution for the market return is more left skewed than the physical distribution, therefore suggesting a positive price of co-kurtosis risk. This is entirely consistent with theory, as explained earlier in this section.

2.3 The General Case

We now examine more general nonlinearities in the SDF. Preference theory is relatively silent about the sign of terms in the SDF beyond the third order, and therefore we do not extend our empirical analysis beyond the cubic SDF. However, while the empirical focus of this paper is on co-skewness and co-kurtosis risk, our approach can be used for any source of risk that is an arbitrary nonlinear (including linear) function of the market return. This does not just include powers of the market return, it includes more complex nonlinear relationships, such as for instance measures of downside risk as in Ang, Chen and Xing (2006). We now present the general result which nests among many other results the results for co-skewness risk in Section 2.1 and co-kurtosis risk in Section 2.2.

Proposition 3 *If the stochastic discount factor (SDF) has the following form:*

$$m_{t+1} = a_t + \sum_k b_{k,t} \left(G_k(R_{m,t+1}) - E_t^P \left[G_k(R_{m,t+1}) \right] \right) + \sum_l c_{l,t} \left(f_{l,t+1} - E_t^P(f_{l,t+1}) \right),$$

then the cross-sectional pricing restrictions are

$$E_t^P(R_{j,t+1}) - R_f = \sum_k \lambda_t^k \beta_{j,t}^k + \sum_l \gamma_t^l \beta_{j,t}^l, \tag{15}$$

and

$$E_t^P(R_{i,t+1}) - R_f = \sum_k \lambda_t^k \beta_{i,t}^k + \sum_l \gamma_t^l \beta_{i,t}^l, \tag{16}$$

where the β_t^k *and* β_t^l *are from the projection of asset returns on* $G_k(R_{m,t+1})$ *and* $f_{l,t+1}$ *respectively, and* γ^l *is the price of risk associated with the factor* f_l*. The price of risk associated with the exposure to a nonlinear function,* G_k*, of the market return,* λ_t^k*, is*

$$\lambda_t^k = E_t^P(G_k(R_{m,t+1})) - E_t^Q(G_k(R_{m,t+1})), \tag{17}$$

where $E_t^P(.)$ *and* $E_t^Q(.)$ *denote the expectation under the physical respectively risk-neutral probability measure.*

Proof. See Appendix A. ■

Proposition 3 shows that the reward for exposure to any nonlinear function G of the market return is determined by the spread between the physical and the risk-neutral expectations of this function. Proposition 3 also shows that the prices of risk associated with the priced factors determine not only the risk premium on the stock but also the spread between the physical expectation of any nonlinear function of the stock returns and its risk-neutral counterpart. Finally, the proposition demonstrates that we can easily incorporate factors that are not necessarily functions of the market return.

3 Estimating the Price of Market Risks

We begin the empirical investigation by documenting the price of co-skewness and co-kurtosis risk using the estimators presented in Propositions 1 and 2. The implementation of our approach requires the estimation of physical and risk-neutral conditional expectations. For the price of co-skewness risk, we need to estimate the second conditional moment under the risk-neutral measure, $E_t^Q(R_{m,t+1}^2)$, and under the physical measure, $E_t^P(R_{m,t+1}^2)$. For the price of co-kurtosis risk, we need to estimate the third conditional moment under the risk-neutral measure $E_t^Q(R_{m,t+1}^3)$ and under the physical measure $E_t^P(R_{m,t+1}^3)$.

3.1 Risk-Neutral Moments from Options

We estimate the risk-neutral moments, which are the inputs in the price of risk formula as shown in propositions 1 and 2, using data on S&P500 index options. This data is collected from OptionMetrics and span the period from January 1996 to December 2012.

The estimation of risk-neutral moments requires a continuum of out-of-the money call and put options which is approximated using cubic spline interpolation techniques. We use the implied volatility estimates reported in OptionMetrics to approximate a continuum of implied volatilities which are in turn converted to a continuum of prices. For strike prices outside the range available, we simply use the implied volatility of the lowest or highest available strike price. See Appendix B for more details.

Following standard practice, we filter out options that (i) violate no-arbitrage conditions, (ii) have missing or extreme implied volatility (larger than 200% or lower than 0.01%), (iii) with open-interest or bid price equal to zero, and (iv) have a bid-ask spread lower than the minimum tick size, i.e., bid-ask spread below \$0.05 for options with prices lower than \$3 and bid-ask spread below \$0.10 for option with prices equal or higher than \$3.

In robustness analysis we use the VIX and SKEW indexes as an alternative estimate for the risk-neutral second and third moments. The data on the VIX and the SKEW index are from the Chicago Board of Options Exchange (CBOE), and are available for the period 1990-2012.

3.2 Physical Moments from Returns

To impose internal consistency between the physical estimates, we want to use estimates of the physical conditional second and third moment that are obtained using the same model. The estimation of conditional higher moments is notoriously difficult. We use a version of the Jondeau and Rockinger (2003) model, who discuss several different implementations. Our implementation is close to the model they refer to as Model 2. This model is among the more parsimonious models they consider. We found it converged well in estimation and for our purposes it is sufficiently richly parameterized. The model is given by

$$R_{m,t} = h_t z_t \qquad\qquad z_t \sim GT(z_t|\eta_t, \lambda_t),$$

where GT denotes the generalized student-t distribution, and where the higher-moment dynamics are modeled via

$$\begin{aligned}
h_t^2 &= a_0 + b_0^+ \left(R_{m,t-1}^+\right)^2 + b_0^- \left(R_{m,t-1}^-\right)^2 + c_0 h_{t-1}^2, \\
\widetilde{\eta}_t &= a_1 + b_1^+ R_{m,t-1}^+ + b_1^- R_{m,t-1}^-, \\
\widetilde{\lambda}_t &= a_2 + b_2^+ R_{m,t-1}^2, \\
\eta_t &= g_{]2,+30]}\left(\widetilde{\eta}_t\right), \text{ and } \lambda_t = g_{]-1,1]}\left(\widetilde{\lambda}_t\right)
\end{aligned}$$

where $R_m^+ = \max(R_m, 0)$ and $R_m^- = \max(-R_m, 0)$. The logistic map

$$g_{]x_L,x_U]}(x) = x_L + \frac{x_U - x_L}{1 + \exp(-x)}$$

ensures that $2 < \eta_t < \infty$ and $-1 < \lambda_t < 1$, which are necessary conditions for the existence of the GT distribution. Note that we have set the conditional mean return to zero here because it is difficult to model and unlikely to matter much for the dynamics of higher moments.

The density of Hansen's (1994) GT distribution is defined by

$$GT(z_t|\eta_t, \lambda_t) = \begin{cases} b_t c_t \left(1 + \frac{1}{\eta_t - 2}\left(\frac{b_t z_t + a_t}{1 - \lambda_t}\right)^2\right)^{-(\eta_t+1)/2} & \text{if } z_t < -a_t/b_t, \\ b_t c_t \left(1 + \frac{1}{\eta_t - 2}\left(\frac{b_t z_t + a_t}{1 + \lambda_t}\right)^2\right)^{-(\eta_t+1)/2} & \text{if } z_t \geq -a_t/b_t, \end{cases}$$

where

$$a_t \equiv 4\lambda_t c_t \frac{\eta_t - 2}{\eta_t - 1}, \quad b_t \equiv 1 + 3\lambda_t^2 - a_t^2, \quad c_t \equiv \frac{\Gamma\left(\left(\eta_t + 1\right)/2\right)}{\sqrt{\pi\left(\eta_t - 2\right)}\Gamma\left(\eta_t/2\right)}.$$

We need the non-centered second and third moments which can be computed as follows

$$E_t^P\left[R_{m,t+1}^2\right] = h_{t+1}^2,$$

and

$$E_t^P\left[R_{m,t+1}^3\right] = h_{t+1}^3\left[m_{3,t+1} - 3a_{t+1}m_{2,t+1} + 2a_{t+1}^3\right]/b_{t+1}^3.$$

where

$$m_{2,t} = 1 + 3\lambda_t^2, \quad m_{3,t} = 16 c_t \lambda_t \left(1 + \lambda_t^2\right)\frac{\left(\eta_t - 2\right)^2}{\left(\eta_t - 1\right)\left(\eta_t - 3\right)},$$

Note that the third moment exists in the model so long as $\eta_t > 3$.

Because estimation of conditional higher moments is difficult, we conducted an extensive robustness analysis. We computed the physical conditional second and third moments from the alternative model in Leon, Rubio and Serna (2005), and the resulting estimates are very similar. For our purpose, both the Jondeau and Rockinger (2003) and Leon, Rubio and Serna (2005) approaches have the advantage that estimates of the conditional third and second moment are internally consistent. If we limit ourselves to focus on co-skewness, then we only require the physical second moment, for which more straightforward estimation techniques are available, such as the standard GARCH model.

3.3 Prices of Co-Skewness and Co-Kurtosis Risk

Armed with the risk-neutral second and third moments from the model-free approach in Section 3.1 and the physical moments from Section 3.2, the estimated price of co-skewness and co-kurtosis risk for month t is now simply

$$\begin{aligned}\widehat{\lambda}_t^{COSK} &= \widehat{E}_t^P(R_{m,t+1}^2) - \widehat{E}_t^Q(R_{m,t+1}^2)\\ \widehat{\lambda}_t^{COKU} &= \widehat{E}_t^P(R_{m,t+1}^3) - \widehat{E}_t^Q(R_{m,t+1}^3)\end{aligned}$$

Table 1 reports descriptive statistics for relevant data and the estimated prices of risk. Figures 1 and 2 depict the time series of the price of co-skewness and co-kurtosis risk with the corresponding estimated physical and risk-neutral moments required to compute these prices. The figures show some spikes surrounding the 1998 LTCM collapse, the WorldCom bankruptcy in 2002, and the subprime crisis. These spikes occur for both the second and third moments, and for risk-neutral as well as physical moments.

The co-skewness risk premium is negative for all months, and the co-kurtosis risk premium is positive for most months. On average the co-skewness price of risk is equal to -0.267, whereas the co-kurtosis price of risk is equal to 0.013 on average. These findings are consistent with theory. Existing empirical studies have also documented negative prices of co-skewness risk and positive prices of co-kurtosis risk. For instance, KL (1976) and HS (2000) find evidence for a negative price of co-skewness risk, while Ang, Chen and Xing (2006) find that stocks with higher co-kurtosis earn higher returns. However, it is critical to emphasize that these existing estimates are typically averages of the price of risk over several years. Most studies estimate prices of risk using a two-pass Fama-MacBeth (1973) setup. Partly because of space constraints, the average estimates of the month-by-month cross-sectional regressions are reported. Even when these averages have the theoretically expected sign, it is possible that the estimates of the price of risk have the opposite sign over extended time periods. What is remarkable about the results reported in Figures 1 and 2 is that the price of risk has the theoretically expected sign in every month in the case of co-skewness risk and almost every month in the case of co-kurtosis risk. Note that there is no guarantee that the results for co-skewness will continue to hold in the future. In fact, in rare instances in the past historical variances have exceeded implied variances. However, we know that such occurrences are rare, therefore all but guaranteeing that we can expect theoretically plausible estimates in the future using this method in the large majority of cases.

4 The Cross-Section of Stock Returns

In this section we present time-series estimates of co-skewness and co-kurtosis betas for different cross-sections of test assets. We then evaluate model fit using an out-of-sample cross-sectional test, and we benchmark the performance of our estimates of co-skewness and co-kurtosis risk against that of existing models.

In our empirical investigation of the cross-section of stock returns, we use several cross-sectional datasets that are commonly used in the existing literature. We use portfolios formed on size and book-to-market ratio, portfolios formed on momentum, and industry portfolios. The data on these portfolios, as well as the data on the Fama-French and momentum factors we use to analyze competing models, are collected from Kenneth French's online data library.

4.1 Portfolio Sorts

Before we evaluate the performance of our newly proposed estimates of the price of co-skewness and co-kurtosis risk in a cross-sectional test, we first provide some more intuition for our results using three very different cross-sections of test assets. This analysis is similar to the one provided by HS (2000) for the 1963-1993 period. Table 2 presents betas estimated using time-series regressions for the following cross-sectional test portfolios: 1) the twenty-five size and book-to-market portfolios of Fama and French (2003); 2) ten momentum portfolios; 3) ten industry portfolios. We report results for two time periods: 1996-2012, which is the sample we obtain when we use OptionMetrics data to compute risk-neutral moments in Section 3.1, and 1990-2012. The latter sample is useful to investigate how robust the patterns in the betas are, and also because we use it in a robustness analysis in Section 5.2 below where we use the CBOE VIX and SKEW as estimates of risk-neutral volatility and skewness.

Panel A of Table 2 reports results for the twenty-five size and book-to-market portfolios of Fama and French (1993). We report the estimate $\beta_{j,t}^{COSK}$ from the projection of portfolio returns on $R_{m,t+1}$, and $R_{m,t+1}^2$, as well as $\beta_{j,t}^{KURT}$ from the projection of portfolio returns on $R_{m,t+1}$, $R_{m,t+1}^2$ and $R_{m,t+1}^3$. For comparison we also report the co-skewness beta β_j^{SKD} in (2) proposed by HS (2000).

For 1996-2012, the excess returns for the different test portfolios clearly indicate a book-to-market effect for smaller stocks, whereas it is less pronounced or non-existing for larger stocks in this sample. Smaller stocks also on average have higher returns than large stocks. Interestingly, part of this cross-sectional variation clearly seems to be captured by co-skewness risk. The correlation between the co-skewness betas $\beta_{j,t}^{COSK}$ and the excess returns is -0.339. Given a theoretically expected negative price of risk, this means that portfolios with greater exposure to systematic skewness have lower returns. This is as expected because higher portfolio skewness is desirable, ceteris paribus. Note that the co-skewness beta β_j^{SKD} from HS (2000) yields a very similar correlation with excess returns of -0.388.

The correlation between the co-kurtosis betas $\beta_{j,t}^{KURT}$ and excess returns is positive, at 0.386, again as expected from theory. It is noteworthy that $\beta_{j,t}^{COKU}$ is indeed extremely large for portfolios consisting of small stocks with high book-to-market, which are the portfolios earning the highest returns.

For the 1990-2012 period, the correlation between the co-skewness betas $\beta_{j,t}^{COSK}$ and excess returns is again negative, at -0.324, and the correlation between the co-kurtosis betas $\beta_{j,t}^{COKU}$ and excess returns is again positive, at 0.354. Moreover, using the alternative co-skewness beta β_j^{SKD} from HS (2000) once again yields a negative correlation at -0.389.

In summary, Panel A of Table 2 clearly indicates that co-skewness is relevant for explaining the variation in excess returns across size and book-to-market portfolios. This finding confirms the results of HS (2000) for the 1963-1993 sample. Moreover, we find that co-kurtosis is relevant for explaining the cross-sectional variation in returns as well. Note of course that the similarity between the co-skewness beta β_j^{SKD} from HS (2000) and our estimate $\beta_{j,t}^{COSK}$ is to be expected, and it says nothing about our estimates of the price of risk. It merely confirms that the measure $\beta_{j,t}^{COSK}$ captures the same cross-sectional patterns as other measures used in the literature.

Panel B repeats this analysis for ten momentum portfolios, and Panel C for ten industry portfolios. Remarkably, the correlation between $\beta_{j,t}^{COSK}$ and excess returns is again negative and economically large for the 1996-2012 sample. Correlations obtained using the alternative co-skewness measure β_j^{SKD} are again rather similar. However, the correlation between the co-kurtosis betas $\beta_{j,t}^{KURT}$ is negative for these data, suggesting that co-kurtosis may not be helpful to explain the cross-sectional variation in returns in this sample.

The large negative correlations associated with $\beta_{j,t}^{COSK}$ in Panel B are very notable. Clearly portfolios with positive momentum, which earn high returns, have much lower exposure to systematic skewness. Perhaps more interestingly, reversing the argument, stocks with negative momentum substantially contribute to portfolio skewness. Ceteris paribus, investors are attracted to lottery-like payoffs in the right tail, and therefore these stocks have low returns in equilibrium. Panel B of Table 2 clearly indicates that co-skewness is a partial explanation of momentum returns. This finding again confirms the findings of HS (2000).

The cross-sectional correlations in Table 2 indicate that co-skewness and, to a lesser extent, co-kurtosis risks help explain the cross-section of stock returns. We now investigate if co-skewness and co-kurtosis risk premia are economically important. Following HS (2000), we find that a one standard deviation decrease (more negative) in the co-skewness beta $\beta_{j,t}^{COSK}$ results in a 3.03% increase in annual expected returns, based on the estimates from the 10 momentum portfolios. Based on the 25 Fama-French portfolios, we obtain a 1.47% increase. HS (2000) document a 2.34% effect of the co-skewness risk on annual expected returns. Note that they use a cross-sectional standard deviation based on individual stocks, as well as a different sample period. We find that the co-kurtosis risk premium is smaller than the co-skewness risk premium. A one standard deviation increase in the co-kurtosis beta $\beta_{j,t}^{COKU}$ leads to an increase of 1.23% in the annual risk premium when we use the 10 momentum portfolios, and an increase of 0.82% for the 25 Fama-French portfolios.

4.2 Fama-MacBeth Estimates of the Price of Co-Skewness and Co-Kurtosis Risk

Table 3 and Figure 3 report estimates of the price of co-skewness and co-kurtosis risk for the 1996-2012 period, obtained using Fama-MacBeth regressions. We use the classical Fama-MacBeth setup, obtaining betas using sixty monthly returns, subsequently running a cross-sectional regression for the next month, and reporting the average of these cross-sectional estimates as well as the t-statistics on these averages. Panel A reports results for the twenty-five size and book-to-market portfolios of Fama and French (2000). The ten momentum portfolios and ten industry portfolios do not offer sufficient cross-sectional variation by themselves, therefore Panel B reports on a combination of the momentum, industry, and Fama-French portfolios.

The CAPM regression in the first column of Table 3 gives an R-square of around 17% in both panels, but the price of market risk λ^{MKT} is estimated with a negative sign in both cases, contrary to theory. The momentum factor and the Fama-French size and book-to-market factors significantly increase the cross-sectional explanatory power, and are estimated with a positive sign, conform with the existing literature. The price of co-skewness risk λ^{COSK} is estimated with the theoretically expected negative sign, confirming the results from Table 2. The estimates are statistically significant at the 5% level in several cases, but not always. The price of co-kurtosis risk λ^{COKU} is estimated with the theoretically expected positive sign, but the estimates are not statistically significant. Co-skewness and co-kurtosis substantially increase the cross-sectional R-square. Although the evidence in Table 2 indicates that co-skewness and co-kurtosis are related to the book-to-market, size, and momentum effects, they do not increase the explanatory power of the cross-sectional regression by as much as these factors.

The explanatory power due to co-skewness and co-kurtosis in these regressions and the estimated signs are of course not our focus. Our interest is in comparing the results from Table 3 with the estimates obtained using our newly proposed method in Table 1. The estimates of λ^{COSK} in Table 3 are all negative and the estimates of λ^{COKU} are all positive. How do they compare to the estimates obtained using our alternative approach in Table 1? Table 1 reports an average price of co-skewness risk of -0.267, larger in absolute value than the estimates in Table 3, which range between -0.073 and -0.153. The estimate of the average price of co-kurtosis risk in Table 1 is 0.013, somewhat larger than the estimates in Table 3.

The advantages of our approach become more obvious when comparing the entire time series of the estimates of the conditional prices of risk in Figure 3, which are estimated using Fama-MacBeth regressions, with our new estimates reported in Figures 1 and 2. Recall that using our proposed method, the estimated price of co-skewness risk is estimated with the theoretically expected negative sign in every month, and the estimated price of co-kurtosis risk is estimated with the theoretically expected positive sign in most months. Note that Table 3 contains the averages and t-statistics for the time series in Figure 3. Clearly, while the price of co-skewness risk is negative on average in this sample, it is positive in many months. Also, the price of co-kurtosis risk is frequently negative. The estimates that do not have the theoretically expected sign are very large in Figure 3 (note that the scales on the axes are different from those in Figures 1 and 2). Moreover, Figure 3 also indicates that at a given point in time, the magnitude and the sign of the estimate of the price of risk may critically depend on the cross-section of portfolios used in estimation.

4.3 Evaluating Model Fit: Out-of-Sample Tests

We now present out-of-sample tests of our new approach and compare the results to existing models. First we motivate the use of out-of-sample tests. Then we discuss the implementation of the tests, and we present the empirical results.

4.3.1 Testing Cross-Sectional Models

To assess the cross-sectional fit of our approach, we benchmark its performance against alternative models. Because our estimation of the price of co-skewness and co-kurtosis risk is very different from the existing literature, it is important to discuss the implementation of tests of cross-sectional asset pricing models in more detail. The literature contains different implementations of tests of cross-sectional models. For instance, an important dimension in which these tests differ is whether the data are averaged first, after which a single cross-sectional regression is run, or whether a cross-sectional regression is run at every time t, after which the estimates of the prices of risk are averaged. The implementation we use in Section 4.2 follows Fama and MacBeth (1973), who run a cross-sectional regression at every time t before averaging the estimates.

Usually these cross-sectional regressions which provide estimates of the prices of risk are also used to evaluate cross-sectional fit and assess the model's performance. For instance, in Table 3 model performance could be judged using the R^2. Even though there are many other related evaluation criteria, in the overwhelming majority of cases these evaluation criteria are similar to the R^2 in Table 3 in the sense that they are in-sample.

It is important to note that in our approach, we construct betas or loadings in exactly the same way as in the traditional Fama-MacBeth setup, but the price of risk is not estimated from a cross-sectional regression. Instead it is estimated as a historical risk premium, and subsequently it is used to assess cross-sectional fit. This difference can best be understood by referring to the well-known case of the CAPM. The CAPM if often evaluated using the Fama-MacBeth approach, by first estimating betas and then running cross-sectional regressions. But alternatively the price of risk for the CAPM could be estimated using the historical market risk premium, and the cross-sectional fit of the CAPM could be evaluated using this price of risk and (the same) estimated betas. It does not make sense to compare the in-sample cross-sectional R-square of the CAPM when the price of risk is estimated in the regression with an R-square obtained by inserting the historical risk premium in the same sample. This amounts to comparing an in-sample fit with an out-of-sample fit. We therefore implement tests of our models using a genuinely out-of-sample approach for all models. Out-of-sample testing of cross-sectional models is becoming increasingly popular, see for instance Simin (2008) and Ferson, Nallareddy, and Xie (2012).

4.3.2 Out-of-Sample Cross-Sectional Tests: Implementation

We present out-of-sample results for two evaluation criteria. Denote the one step-ahead forecast provided by the model and by $\widehat{R}_{j,t+1}^{Model}$. In our implementation, which is recursive, this forecast uses information available up to time t. The first evaluation criterion is the mean of the squared forecast error, also used by Simin (2008), which is given by

$$RMSFE_{OS} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(R_{j,t} - \widehat{R}_{j,t+1}^{Model}\right)^2} \qquad (18)$$

where T is the number of time periods in the sample. We report this measure for each individual portfolio j. Our second evaluation criteria is adapted from the time-series literature. We use the out of sample R-square suggested by Campbell and Thompson (2008), which has become the standard in the time-series literature, see for instance Rapach and Zhou (2013). The out-of-sample R_{OS}^2 is defined by

$$R_{OS}^2 = 1 - \frac{\sum_t \left(R_{j,t} - \widehat{R}_{j,t+1}^{Model}\right)^2}{\sum_t \left(R_{j,t} - \overline{R}_{j,t-59:t}\right)^2} \qquad (19)$$

where $\overline{R}_{j,t-59:t} = \frac{1}{60}\sum_{s=t-59}^{t} R_{j,s}$. This R-square can again be computed for every portfolio, but because of space constraints we report the average across portfolios for each model.

Note that this out-of-sample R-square uses the historical return on the test portfolio as a benchmark. If a candidate model performs as well as the historical return on the test portfolio, the resulting R-square will be zero. R-squares will be negative for models that do not perform well in forecasting. Consequently, the values of this out-of-sample R-square should not be confused with the R-squares one typically obtains from a cross-sectional or time-series regression, for example. In fact, R-squares can be expected to be very small, and a small positive R-square is an indicator of success. See Campbell and Thompson (2008), Rapach, Strauss, and Zhou (2010), and Rapach and Zhou (2013) for a detailed discussion.

We compare the cross-sectional performance of our newly proposed estimates of the price of co-skewness and co-kurtosis risk to a number of other specifications based on these two evaluation criteria. One set of specifications is based on historical risk premia, in the other one the risk premia are estimated using cross-sectional regressions. The models that use cross-sectional regressions to estimate the risk premia are: the CAPM, the Fama-French three-factor model (FF), and the CAPM augmented with co-skewness and co-kurtosis. These two models are denoted by CAPM+COSK and CAPM+ COKU respectively. For comparison we also include the CAPM augmented with co-skewness as measured by β_j^{SKD} in (2). The specifications based on historical risk premia are: CAPM, COSK, COKU, CAPM+COSK, and CAPM+COKU. To provide more intuition, we explain the implementation of the two types of specifications using the CAPM as an example.

For the CAPM, the one step-ahead forecast of $\widehat{R}_{j,t+1}^{CAPM}$ using information available up to time t is

$$\widehat{R}_{j,t+1}^{CAPM} = \widehat{\lambda}_t^{mkt} \widehat{\beta}_{j,t}^{mkt} \tag{20}$$

The betas for both implementations are the same, and are obtained by regressing R_j on R_m, using a rolling window of 60 months from $t-59$ to t. However, estimates of the covariance price of risk, $\widehat{\lambda}_t^{mkt}$, are obtained in two ways. The first approach uses the sample mean of the market excess return over the past 60 months. The second approach is to estimate the price of risk using a cross-sectional regression:

$$R_{j,t} = \boldsymbol{\lambda}_t^{mkt} \, \widehat{\boldsymbol{\beta}}_{j,t-1}^{mkt} + u_{j,t}, \quad j = 1, ...N \tag{21}$$

Note that in principle we can at each time t use this price of risk $\boldsymbol{\lambda}_t$ to construct the forecast of $\widehat{R}_{j,t+1}^{CAPM}$. However, we found that this leads to extremely poor forecasts, which is due to the time variation in these cross-sectional estimates, as evidenced by the estimates for co-skewness and co-kurtosis in Figure 3. To provide better out-of-sample competitors for our estimators of co-skewness and co-kurtosis that use historical risk premia, we therefore use averages of the cross-sectional averages of $\boldsymbol{\lambda}_t$ for the past 60 months, which provided better forecasts.

4.3.3 Out-of-Sample Cross-Sectional Tests: Empirical Results

We present results for three sets of test portfolios: 1) ten industry portfolios; 2) the 25 size and book-to-market portfolios from Fama and French (1993); and 3) ten momentum portfolios.

Table 4 presents the results for the out-of-sample RMSFEs, which we have multiplied by 100, following the convention adopted by Simin (2008). To interpret these numbers, note that if the forecast is the historical average, the magnitude should be similar to a monthly volatility. For a stock with 30% annual volatility, the monthly volatility is 8.660%. The second and third columns present results that are obtained using our newly proposed estimates of the price of co-skewness and co-kurtosis risk. The co-skewness based forecasts, in column 2, provide the lowest forecast errors, 5.995 on average. The co-kurtosis based forecasts also provide a lower forecast error, 6.024, than all forecast errors based on cross-sectional regressions in columns 6-10. Note also that the forecast errors based on co-skewness and co-kurtosis are also smaller than the forecast error for the $CAPM$ forecast error based on the historical market risk premium, which is 6.083. Moreover, combining the market risk premium with the co-skewness and co-kurtosis risk premium increases the forecast error.

The out-of-sample RMSFEs provide a useful ranking of the models, but it takes some effort to interpret the magnitudes. The out-of-sample R-square R^2_{OS} evaluation criterion is perhaps easier to understand intuitively. Table 5 presents the results, which for each model are averaged over all portfolios. Recall that a positive out-of-sample R-square means that the model forecasts better than the historical average return on the asset. The first column in each panel lists the results using the entire 1996-2012 sample. We also report results for the first and the second half of the sample, and for all years with cumulative positive and negative returns, because in our sample all models do a lot better in years with negative cumulative returns.

We report results for the same ten specifications reported in Table 4, five based on cross-sectional regressions and five based on historical risk premiums. The performance of our newly proposed co-skewness measure $COSK$ in the second row of each panel is impressive. It yields a positive R-square twelve times out of fifteen. The out-of-sample performance of the co-kurtosis model $COKU$ is mixed, with eight positive R-squares out of fifteen. In comparison, the $CAPM$ yields six positive R-squares when historical risk premiums are used. Results are rather disappointing when combining $COSK$ with the $CAPM$ in row 4, and especially when combining $COKU$ with the $CAPM$ in row 5.

It is worth remembering that in a genuine out-of-sample setting, these very small positive R-squares are economically meaningful. This criterion is typically used in the time-series literature, and even there R-squares of 1-2% are the exception rather than the rule, with many candidate forecasts yielding negative R-squares, see Campbell and Thompson (2008), Rapach, Strauss, and Zhou (2010), and Rapach and Zhou (2013). The performance of the newly proposed estimate of the price of co-skewness risk is therefore impressive, especially because forecasting with a cross-sectional model is even harder than time series forecasting. This is also emphasized by a more detailed inspection of the performance of the benchmark models implemented using cross-sectional regressions in rows 6-10. The $CAPM$ yields eight positive R-squares when cross-sectional regressions are used (row 6 of each panel). Results are similar for the $CAPM + COSK$ and $CAPM + SKD$ models. The Fama-French model yields negative R-squares in many cases. It may seem surprising that the FF model performs so poorly for the case of the 25 size and book-to-market portfolios in Panel B, but note that the FF model is not typically evaluated in a genuine out-of-sample setting.

5 Robustness

We now consider several robustness exercises. First we consider alternative measures of conditional physical second and third moments. Subsequently we measure the second risk-neutral moment by the VIX and the third risk-neutral moment by the SKEW, as an alternative to the use of the model-free method of Bakshi and Madan (2000) and Bakshi, Kapadia, and Madan (2003).

5.1 Alternative Physical Moments

We first investigate the robustness of our results to different measurements of the physical moments. The estimates of the price of co-skewness and co-kurtosis risk in Figures 1-2 and Table 1 use the model of Jondeau and Rockinger (2003). We obtained similar results using the method of Leon et al. (2005). However, estimating conditional higher moments is notoriously difficult, and it is possible that the estimation approach affects our estimates of the price of risk. We therefore also investigate a dramatically different approach to estimating the physical moments which imposes far less structure. Specifically, we construct monthly measures of the conditional second and third moment by using realized measures based on daily data, and subsequently imposing an autoregressive structure. In other words, we use

$$\kappa_t^2 = a_0 + a_1 \kappa_{t-1}^2 + u_t^3 \quad (22)$$

$$\kappa_t^3 = b_0 + b_1 \kappa_{t-1}^3 + u_t \quad (23)$$

where $\kappa_t^2 = \sum_{d \epsilon t} R_{d,t}^2$, $\kappa_t^3 = \sum_{d \epsilon t} R_{d,t}^3$, and $R_{d,t}$ is the daily return in day d of month t.

Table 6 presents the results. To save space, we limit ourselves to the out-of-sample R-squares R_{OS}^2. The R-squares are positive for the $COSK$ model, except for years with positive returns, and the magnitudes of the R-squares are similar to those in Table 5. The $COKU$ model does not perform as well. As in Table 5, combining $COSK$ with other factors leads to a decrease in model performance.

5.2 Alternative Risk-Neutral Moments

We now investigate the robustness of the results to an alternative computation of the second risk-neutral moment. Rather than using the model-free approach in Bakshi and Madan (2000) and Bakshi, Kapadia and Madan (2003), we simply use the VIX as an estimate of the risk-free second moment and the SKEW as a measure of the risk-free third moment. Both time series are available from the CBOE for the period 1990-2012. This approach has a number of advantages. The construction of the VIX and the SKEW is exogenous to our experiment, and so it is not possible to design it to maximize performance. Even more importantly, VIX and SKEW are available for a much longer sample period than the model-free second and third moment. For studies that use the VIX as a proxy for the risk-neutral second moment see, for instance, Bollerslev Tauchen and Zhou (2009).

Table 7 presents the results. We present results for the entire sample, for the first and second halves of the sample, and for years with negative and positive cumulative returns. The most important conclusion is that the results for $COSK$ in the second row of each panel are again good compared to the other models. Results for $COKU$ in row three are not as good. Note also that the performance of the models that use cross-sectional regression are worse than in Table 5, and in some cases these models perform very poorly. It seems that the 1990-2012 time period is more challenging for the out-of-sample performance of all models.

6 Conclusion

We propose an alternative strategy to estimate the price of possibly nonlinear exposures to market risk, which avoids the errors inherent in the cross-sectional regression approach. The key difference between our approach and existing studies is that we explicitly impose the resulting pricing restrictions on both stocks and contingent claims. We study two important applications of our general approach. First, the price of co-skewness risk in our framework corresponds to the spread between the physical and the risk-neutral second moment. Second, the price of co-kurtosis risk is similarly given by the spread between the physical and the risk-neutral third moment.

Using monthly data for the period 1996-2012, we find that the price of co-skewness risk has the theoretically expected negative sign in each month, and the price of co-kurtosis risk has the theoretically expected positive sign in most months. While the prices of risk obtained using the Fama-MacBeth approach have the theoretically anticipated signs on average, they often yield very large estimates with a sign that contradicts economic theory, in contrast with our approach. An out-of-sample analysis of factor models with co-skewness and co-kurtosis risk indicates that the new estimates of the price of risk improve the models' performance. The models also robustly outperform competitors such as the CAPM and the Fama-French model.

Some questions remain, and a number of extensions could prove interesting. First, while the estimated price of co-skewness risk leads to a more than satisfactory out-of-sample cross-sectional fit when used by itself, its performance is worse when combined with the CAPM risk factor. It may prove useful to further investigate the resulting biases. Second, the differential performance of co-skewness in down and up markets deserves further study. Third, the focus of this paper is on improving measurement. While we believe that our measure of the price of co-skewness risk improves on existing techniques, we worry that the betas we use in the analysis may be subject to substantial biases. Improved estimation of betas may be worth exploring, and may lead to better out-of-sample performance. The estimation approach proposed by Bali and Engle (2010) may be especially promising in this regard.

Appendix A: Proof of Propositions 1, 2 and 3

Linear factor models, in which the stochastic discount factor is $m_{t+1} = a_t + \mathbf{b}_t' \left(\widetilde{\mathbf{f}}_{t+1} - E_t^P(\widetilde{\mathbf{f}}_{t+1}) \right) = a_t + \mathbf{b}_t'\mathbf{f}_{t+1}$, are equivalent to beta-representation models with the vector of risk factors $\mathbf{f}$

$$E_t^P(R_{j,t+1}) - R_{f,t} = \boldsymbol{\lambda}_t' \boldsymbol{\beta}_{j,t}, \tag{24}$$

where $\boldsymbol{\lambda}_t' = \frac{-1}{a_t}\mathbf{b}_t' E_t^P(\mathbf{f}_{t+1}\mathbf{f}_{t+1}')$, $(R_{f,t}+1) = \frac{1}{a_t} = \frac{1}{E_t^P(m_{t+1})}$, $\boldsymbol{\beta}_{j,t} = \left[E_t^P(\mathbf{f}_{t+1}\mathbf{f}_{t+1}')\right]^{-1} E_t^P(\mathbf{f}_{t+1}R_{j,t+1})$, and $R_{j,t+1}$ is the return on a stock j (see for instance Cochrane (2005)). Since the pricing kernel prices all the assets including contingent claims, the above equation also holds for any claim i whose price is contingent on the stock j and has a payoff function $\Psi(R_{j,t+1})$, for any function $\Psi(.)$. From equation (24) we have

$$E_t^P \left(\frac{\Psi(R_{j,t+1}) - P_{i,t}}{P_{i,t}} \right) - R_{f,t} = \boldsymbol{\lambda}_t' \boldsymbol{\beta}_{i,t}, \tag{25}$$

where $P_{i,t}$ is the price of the contingent claim i. Using the definition of $\boldsymbol{\beta}_{i,t}$ we have

$$E_t^P \left(\frac{\Psi(R_{j,t+1}) - P_{i,t}}{P_{i,t}} \right) - R_{f,t} = \boldsymbol{\lambda}_t' \left[E_t^P((\mathbf{f}_{t+1}\mathbf{f}_{t+1}')\right]^{-1} E_t^P \left((\mathbf{f}_{t+1} \frac{\Psi(R_{j,t+1}) - P_{i,t}}{P_{i,t}} \right). \tag{26}$$

Rearranging and using $E_t^P(\mathbf{f}_{t+1}) = 0$ gives

$$E_t^P \left[\Psi(R_{j,t+1})\right] - P_{i,t}\left(1 + R_{f,t}\right) = \boldsymbol{\lambda}_t' \left[E_t^P((\mathbf{f}_{t+1}\mathbf{f}_{t+1}')\right]^{-1} E_t^P \left[(\mathbf{f}_{t+1}\Psi(R_{j,t+1})\right] \tag{27}$$

The no-arbitrage condition ensures the existence of at least one risk-neutral measure Q such that $P_{i,t} = \frac{1}{\left(1+R_{f,t}\right)} E_t^Q \left[\Psi(R_{j,t+1})\right]$. Therefore, we obtain

$$E_t^P \left[\Psi(R_{j,t+1})\right] - E_t^Q \left[\Psi(R_{j,t+1})\right] = \boldsymbol{\lambda}_t' \boldsymbol{\beta}_{\Psi,t} \tag{28}$$

where $\boldsymbol{\beta}_{\Psi,t}$ is from the projection of $\Psi(R_{j,t+1})$ on $\mathbf{f}_{t+1}$.

Assume that $m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right)$, then applying equation (28) for $\Psi = R_m$ (i.e., $\boldsymbol{\beta}_{\Psi,t} = \begin{bmatrix} 1 & 0 \end{bmatrix}'$) we recover equation (5) of Proposition 1, and applying equation (28) for $\Psi = R_m^2$ (i.e., $\begin{bmatrix} 0 & 1 \end{bmatrix}'$) we recover equation (6) of the same proposition.

Following the same logic, we can establish propositions 2 and 3. Assume that $m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right) + b_{3,t}\left(R_{m,t+1}^3 - E_t^P(R_{m,t+1}^3)\right)$, then, as previously, applying equation (28) for $\Psi = R_m$ (i.e., $\boldsymbol{\beta}_{\Psi,t} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}'$) we recover equation (12) of proposition 2, and applying equation (28) for $\Psi = R_m^2$ (i.e., $\boldsymbol{\beta}_{\Psi,t} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}'$), we recover equation (13) of the same proposition. In addition, applying equation (28) for $\Psi = R_m^3$ (i.e., $\boldsymbol{\beta}_{\Psi,t} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}'$), we obtain equation (14) of proposition 2.

Finally, if $m_{t+1} = a_t + \sum_k b_{k,t} \left(G_k(R_{m,t+1}) - E_t^P \left[G_k(R_{m,t+1}) \right] \right) + \sum_l c_{l,t} \left(f_{l,t+1} - E_t^P(f_{l,t+1}) \right)$, then applying equation (28) for $\Psi = G_k(R_{m,t+1})$ (i.e., $\boldsymbol{\beta}_{\Psi,t}$ is a unit vector with the k^{th} element equal to one and all the other elements equal to zero), we obtain obtain equation (17) of Proposition 3.

Appendix B: Extracting Option Implied Moments

Bakshi and Madan (2000) have shown that any twice-continuously differentiable payoff function, $H[S]$, can be spanned by a portfolio of risk-free bonds, the underlying asset and out-of-the-money calls and puts as follows

$$\begin{aligned} H[S] = H\left[\overline{S}\right] + \left(S - \overline{S}\right) H_S\left[\overline{S}\right] + \int_{\overline{S}}^{\infty} H_{SS}[K] (S-K)^+ dK \\ + \int_0^{\overline{S}} H_{SS}[K] (K-S)^+ dK. \end{aligned} \quad (29)$$

The prices of these contracts are

$$\begin{aligned} E_t^Q \left\{ e^{-r\tau} H[S] \right\} = \left(H\left[\overline{S}\right] - \overline{S} H_S\left[\overline{S}\right] \right) e^{-r\tau} + H_S\left[\overline{S}\right] S(t) \\ + \int_{\overline{S}}^{\infty} H_{SS}[K]\, C(t,\tau;K)\, dK + \int_0^{\overline{S}} H_{SS}[K]\, P(t,\tau;K)\, dK. \end{aligned} \quad (30)$$

where $C_t(\tau, K)$ and $P_t(\tau, K)$ are prices of the European call and put options with time-to-maturity τ and strike price K. As a result, we can calculate the prices of derivatives whose payoffs only depend on the future S, given the prices of (i) the risk free zero coupon bond, r, (ii) the current value of the underlying stock, $\overline{S}$, and (iii) a series of OTM calls and puts.

For our purposes, let $R(t,\tau) = \ln S(t+\tau) - \ln S(t)$, and first consider the function

$$H[S_{t+\tau}] = R_{t+\tau}^2 = (\ln S_{t+\tau} - \ln S_t)^2 \quad (31)$$

Using this, we can get the risk-neutral raw second moment via

$$\begin{aligned} E_t^Q\left[R_{t+\tau}^2\right] &= e^{r\tau}\int_{S_t}^{\infty}\frac{2\left(1-\ln\left[K/S_t\right]\right)}{K^2}C_t\left(\tau,K\right)dK \\ &\quad +e^{r\tau}\int_0^{S_t}\frac{2\left(1+\ln\left[S_t/K\right]\right)}{K^2}P_t\left(\tau,K\right)dK. \end{aligned}$$

Now, let

$$H[S_{t+\tau}] = R_{t+\tau}^3 = \left(\ln S_{t+\tau} - \ln S_t\right)^3 \tag{32}$$

then we get the option-implied raw third moment via

$$\begin{aligned} E_t^Q\left[R_{t+\tau}^3\right] &= e^{r\tau}\int_{S_t}^{\infty}\frac{6\ln\left[K/S_t\right]-3\left(\ln\left[K/S_t\right]\right)^2}{K^2}C_t\left(\tau,K\right)dK \\ &\quad -e^{r\tau}\int_0^{S_t}\frac{6\ln\left[S_t/K\right]+3\left(\ln\left[S_t/K\right]\right)^2}{K^2}P_t\left(\tau,K\right)dK. \end{aligned}$$

When computing these moments, we eliminate put options with strike prices of more than 105% of the underlying asset price ($K/S > 1.05$) and call options with strike prices of less than 95% of the underlying asset price ($K/S < 0.95$). We only estimate the moments for days that have at least two OTM call prices and two OTM put prices available.

Since we do not have a continuity of strike prices, we calculate the integrals using cubic splines. For each maturity, we interpolate implied volatilities using a cubic spline across moneyness levels (K/S) to obtain a continuum of implied volatilities. For moneyness levels below or above the available moneyness level in the market, we use the implied volatility of the lowest or highest available strike price. After implementing this interpolation-extrapolation technique, we obtain a fine grid of implied volatilities for moneyness levels between 1% and 300%. We then convert these implied volatilities into call and put prices using the following rule: moneyness levels smaller than 100% ($K/S < 1$) are used to generate put prices and moneyness levels larger than 100% ($K/S > 1$) are used to generate call prices using trapezoidal numerical integration. Linear interpolation between maturities is used to calculate the moments for a fixed 30-day horizon.

References

[1] Ang, A., J. Chen, and Y. Xing (2006), Downside Risk, *Review of Financial Studies*, 19, 1191–1239.

[2] Arditti, F. (1967), Risk and the Required Returns on Equity, *Journal of Finance*, 22, 19–36.

[3] Bakshi, G., N. Kapadia, and D. Madan (2003), Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options, *Review of Financial Studies*, 16 (1), 101–143.

[4] Bakshi, G., and D. Madan (2000), Spanning and Derivative Security Valuation, *Journal of Financial Economics*, 55, 205–238.

[5] Bakshi, G., and D. Madan (2006), A Theory of Volatility Spreads, *Management Science*, 52, 1945–1956.

[6] Bakshi, G., D. Madan and G. Panayotov, (2010), Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels, *Journal of Financial Economics*, 97, 130–154.

[7] Bali, T. G., O. Demirtas, and H. Levy (2009), Is There an Intertemporal Relation between Downside Risk and Expected Returns? *Journal of Financial and Quantitative Analysis*, 44, 883–909.

[8] Bali, T. G., and R. F. Engle (2010), The Intertemporal Capital Asset Pricing Model with Dynamic Conditional Correlations, *Journal of Monetary Economics*, 57, 377–390.

[9] Bansal, R., and S. Viswanathan (1993), No-Arbitrage and Arbitrage Pricing: A New Approach, *Journal of Finance*, 48, 1231–1262.

[10] Black, F., and M. Scholes (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637–654.

[11] Bondarenko, O. (2003), Why Are Put Options So Expensive? Working Paper, University of Illinois, Chicago.

[12] Bollerslev, T., G. Tauchen, and H. Zhou (2009), Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies*, 22, 4463–4492.

[13] Broadie, M., M. Chernov, and M. Johannes (2009), Understanding Index Option Returns, *Review of Financial Studies*, 22 (11), 4493–4529.

[14] Campbell, J. Y., and S. B. Thompson (2008), Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies*, 21, 1509–31.

[15] Carhart, M. (1997), On Persistence of Mutual Fund Performance, *Journal of Finance*, 52, 57–82.

[16] Carr, P., and L. Wu (2009), Variance Risk Premia, *Review of Financial Studies*, 22, 1311–1341.

[17] Chabi-Yo, F. (2008), Conditioning Information and Variance Bounds on Pricing Kernels with Higher-Order Moments: Theory and Evidence, *Review of Financial Studies*, 21, 181–231.

[18] Chan, L. K.C., J. Karceski, and J. Lakonishok (2003), The Level and Persistence of Growth Rates, *Journal of Finance*, 58, 643–684.

[19] Cochrane, J. (2005), *Asset Pricing*, Princeton University Press, Princeton, NJ.

[20] Coval, J. D., and T. Shumway (2001), Expected Option Returns, *Journal of Finance*, 56, 983–1009.

[21] Dittmar, R. F. (2002), Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns, *Journal of Finance*, 57, 369–403.

[22] Driessen, J. , P. J. Maenhout, and G. Vilkov (2009), The Price of Correlation Risk: Evidence from Equity Options, *Journal of Finance*, 64, 1377–1406.

[23] Fama, E. F., and K. R. French (1993), Common Risk Factors in the Returns on Stocks and Bonds, J*ournal of Financial Economics*, 33, 3–56.

[24] Fama, E. F., and J. D. MacBeth (1973), Risk, Return and Equilibrium: Empirical Tests, *Journal of Political Economy*, 81, 607–636.

[25] Ferson, W., S. Nallareddy, and B. Xie (2012), The "Out-of-Sample" Performance of Long Run Risk Models, *Journal of Financial Economics*, 107, 537–556.

[26] Golec, J., and M. Tamarkin (1998), Bettors Love Skewness, Not Risk, at the Horse Track, *Journal of Political Economy*, 106, 205–225.

[27] Guidolin, M., and A. Timmermann (2008), International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences, *Review of Financial Studies*, 21, 889–935.

[28] Hansen, B. (1994), Autoregressive Conditional Density Estimation, *International Economic Review*, 35, 705-730.

[29] Harvey, C. R., and A. Siddique (2000), Conditional Skewness in Asset Pricing, *Journal of Finance*, 55 (3), 1263–1295.

[30] Jackwerth, J., and M. Rubinstein (1996), Recovering Probability Distributions from Option Prices, *Journal of Finance* 51, 1611–1631.

[31] Jondeau, E., and M. Rockinger (2003), Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements, *Journal of Economic Dynamics and Control*, 27, 1699–1737.

[32] Kraus, A., and R. H. Litzenberger (1976), Skewness Preference and the Valuation of Risk Assets, *Journal of Finance*, 31, 1085–1100.

[33] Leland, H. (1997), Beyond Mean-Variance: Performance Measurement of Portfolios Using Options or Dynamic Strategies, Working paper, University of California, Berkeley.

[34] León, A., G. Rubio, and G. Serna (2005), Autoregressive Conditional Volatility, Skewness and Kurtosis, *The Quarterly Review of Economics and Finance*, 45, 599–618.

[35] Lewellen, J., S. Nagel, and J. Shanken (2010), A Skeptical Appraisal of Asset Pricing Tests, *Journal of Financial Economics*, 96, 175–194.

[36] Liew, J., and M. Vassalou (2000), Can Book-to-Market, Size, and Momentum Be Risk Factors that Predict Economic Growth?, *Journal of Financial Economics*, 57, 221–245.

[37] Lim, K.-G. (1989), A New Test of the Three-Moment Capital Asset Pricing Model, *Journal of Financial and Quantitative Analysis,* 24, 205-216.

[38] Rapach, D. E., J. K. Strauss and G. Zhou, (2010), Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, *Review of Financial Studies*, 23, 821–862.

[39] Rapach, D. E., and G. Zhou (2013), Forecasting Stock Returns, in the *Handbook of Economic Forecasting*, Volume 2, edited by G. Elliott and A. Timmermann, Elsevier: San-Diego, pp. 328–383.

[40] Rubinstein, M. (1973), The Fundamental Theory of Parameter-Preference Security Valuation, *Journal of Financial and Quantitative Analysis*, 8, 61–69.

[41] Santa-Clara, P., and A. Saretto (2009), Option Strategies: Good Deals and Margin Calls, *Journal of Financial Markets*, 12, 391–417.

[42] Simin, T. (2008), The Poor Predictive Performance of Asset Pricing Models, *Journal of Financial and Quantitative Analysis*, 43, 355-380.

[43] Scott, R. C., and P. A. Horvath (1980), On the Direction of Preference for Moments of Higher Order than the Variance, *Journal of Finance*, 35, 915–919.

Figure 1: Time-Varying Price of Coskewness Risk

We plot the times series for the conditional physical and risk-neutral second moments (% per month), and the price of co-skewness risk. The physical second moment is estimated using the autoregressive conditional volatility, skewness, and kurtosis model of Jondeau and Rockinger (2003), in which market return follows

$$R_{m,t} = h_t z_t, \qquad\qquad z_t \sim GT(z_t|\eta_t, \lambda_t),$$

where GT denotes the generalized Student t distribution. The moment dynamics are driven by

$$\begin{aligned} h_t^2 &= a_0 + b_0^+ \left(R_{m,t-1}^+\right)^2 + b_0^- \left(R_{m,t-1}^-\right)^2 + c_0 h_{t-1}^2, \\ \eta_t &= g_\eta \left(a_1 + b_1^+ R_{m,t-1}^+ + b_1^- R_{m,t-1}^-\right), \\ \lambda_t &= g_\lambda \left(a_2 + b_2^+ R_{m,t-1}^2\right), \end{aligned}$$

where g_η and g_λ denote logistic transformations provided in Section 3.2. The risk-neutral second moment is estimated using the model-free approach in Bakshi and Madan (2000) and Bakshi, Kapadia and Madan (2003). The time-varying price of co-skewness risk is equal to the spread between the physical and risk-neutral moments. The sample period is from January 1996 through December 2012.

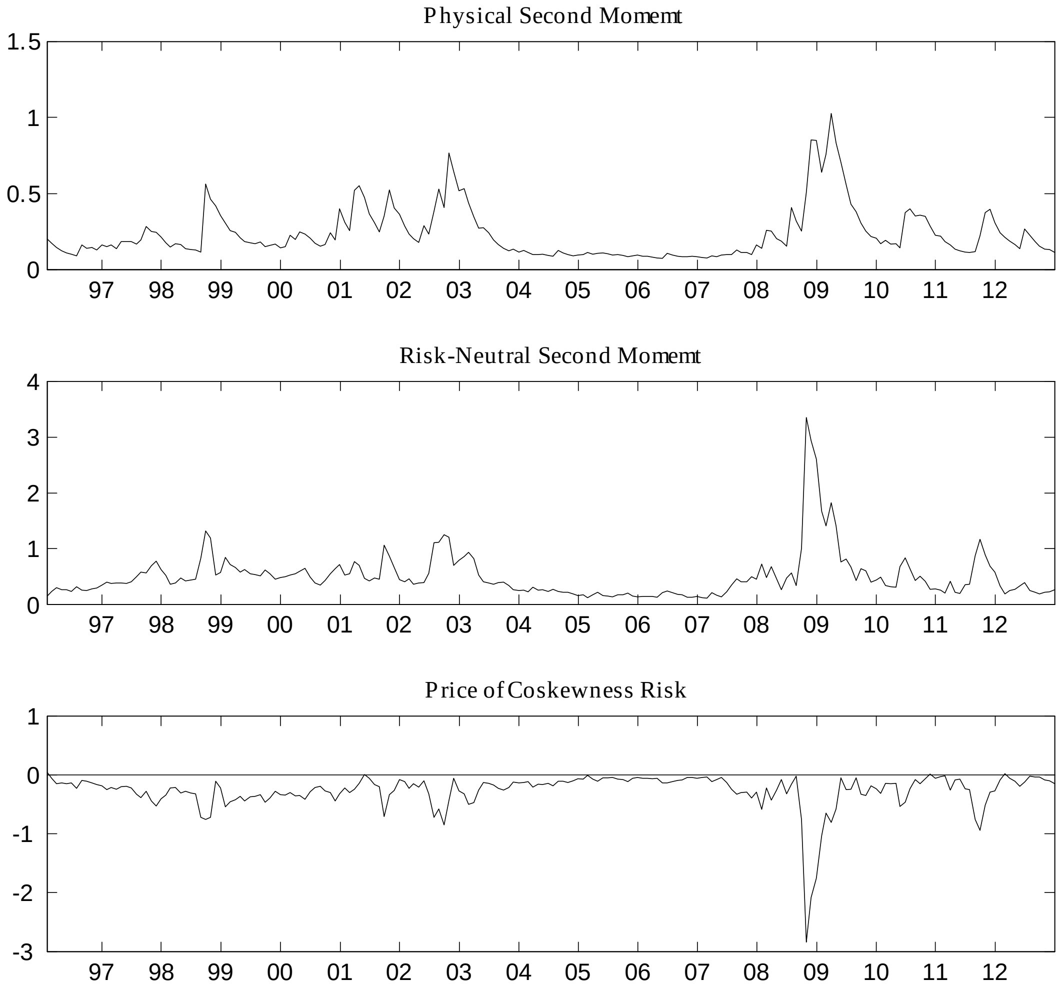

Figure 2: Time-Varying Price of Cokurtosis Risk

We plot the times series for the conditional physical and risk-neutral third moments (% per month). We also plot the price of co-kurtosis risk. The physical third moment is estimated using the autoregressive conditional volatility, skewness, and kurtosis model of Jondeau and Rockinger (2003) in which the market return follows

$$R_{m,t} = h_t z_t, \qquad z_t \sim GT(z_t|\eta_t, \lambda_t),$$

where GT denotes the generalized Student t distribution. The moment dynamics are driven by

$$\begin{aligned} h_t^2 &= a_0 + b_0^+ \left(R_{m,t-1}^+\right)^2 + b_0^- \left(R_{m,t-1}^-\right)^2 + c_0 h_{t-1}^2, \\ \eta_t &= g_\eta \left(a_1 + b_1^+ R_{m,t-1}^+ + b_1^- R_{m,t-1}^-\right), \\ \lambda_t &= g_\lambda \left(a_2 + b_2^+ R_{m,t-1}^2\right), \end{aligned}$$

where g_η and g_λ denote logistic transformations provided in Section 3.2. The risk-neutral third moment is estimated using the model-free approach in Bakshi and Madan (2000) and Bakshi, Kapadia and Madan (2003). The time-varying price of co-kurtosis risk is equal to the spread between the physical and risk-neutral moments. The sample period is from January 1996 through December 2012

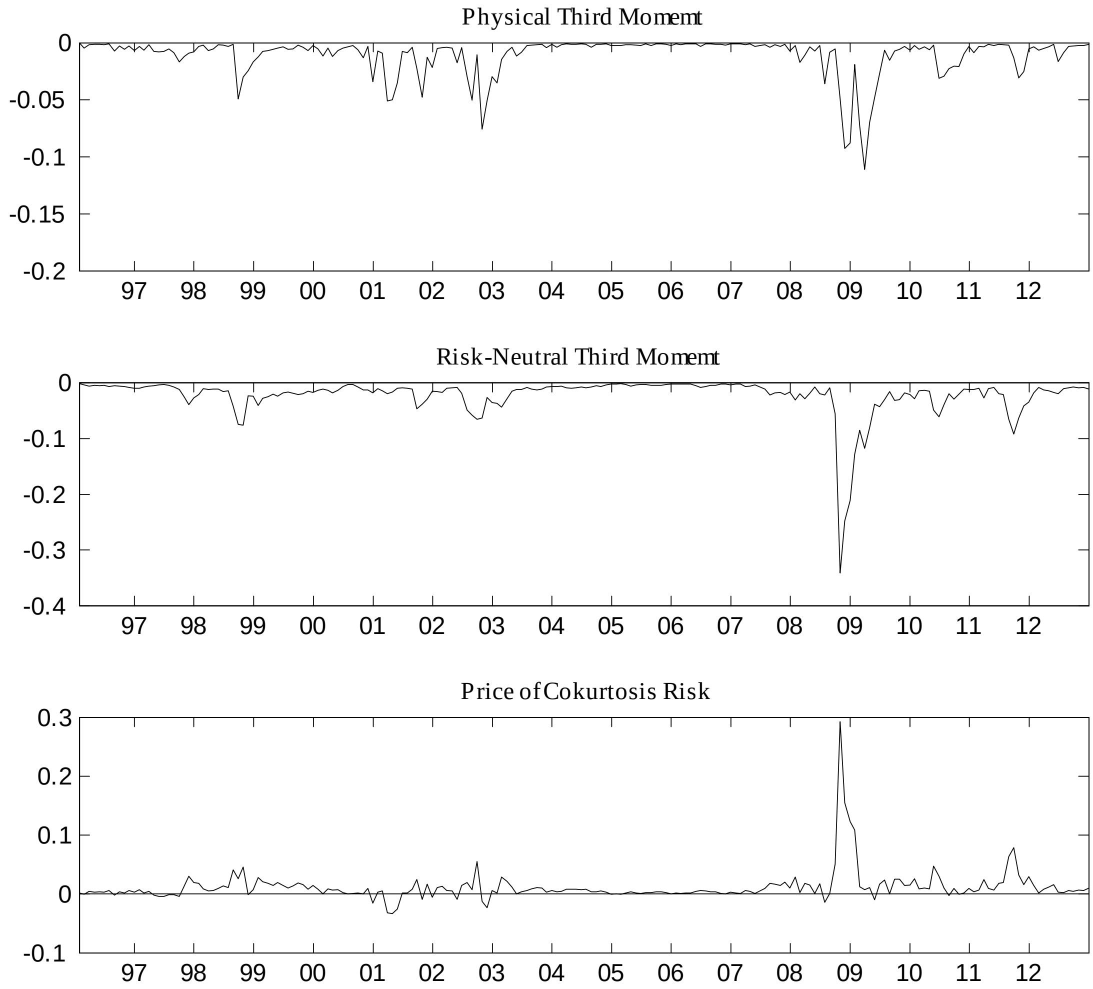

Figure 3: Cross-Sectional Prices of Risk

We plot the cross-sectional prices of co-skewness and co-kurtosis risk over time. Each month, we estimate the co-skewness beta using a 60-month rolling window of monthly returns from the following time series regression

$$R_{j,t} - r_f = \alpha_j + \beta_{j,t}^{MKT} R_{MKT,t} + \beta_{j,t}^{COSK} R_{MKT,t}^2 + \varepsilon_{j,t}.$$

We then run the following cross-sectional regression using returns over the next month

$$R_{j,t+1} - r_f = \lambda_{0,t} + \beta_{j,t}^{MKT} \lambda_{MKT,t} + \beta_{j,t}^{COSK} \lambda_{COSK,t} + e_{j,t}.$$

For the price of co-kurtosis risk we proceed in a similar way. First, we run a time series regression

$$R_{j,t} - r_f = \alpha_j + \beta_{j,t}^{MKT} R_{MKT,t} + \beta_{j,t}^{COKU} R_{MKT,t}^3 + \varepsilon_{j,t},$$

and then we run a cross-sectional regression

$$R_{j,t+1} - r_f = \lambda_{0,t} + \beta_{j,t}^{MKT} \lambda_{MKT,t} + \beta_{j,t}^{COKU} \lambda_{COKU,t} + e_{j,t}.$$

We consider two sets of portfolios: 25 Size/Book-to-Market (top panels) and 25 Size/Book-to-Market + 10 Momentum + 10 Industry (bottom panels). The sample period is from January 1996 through December 2012.

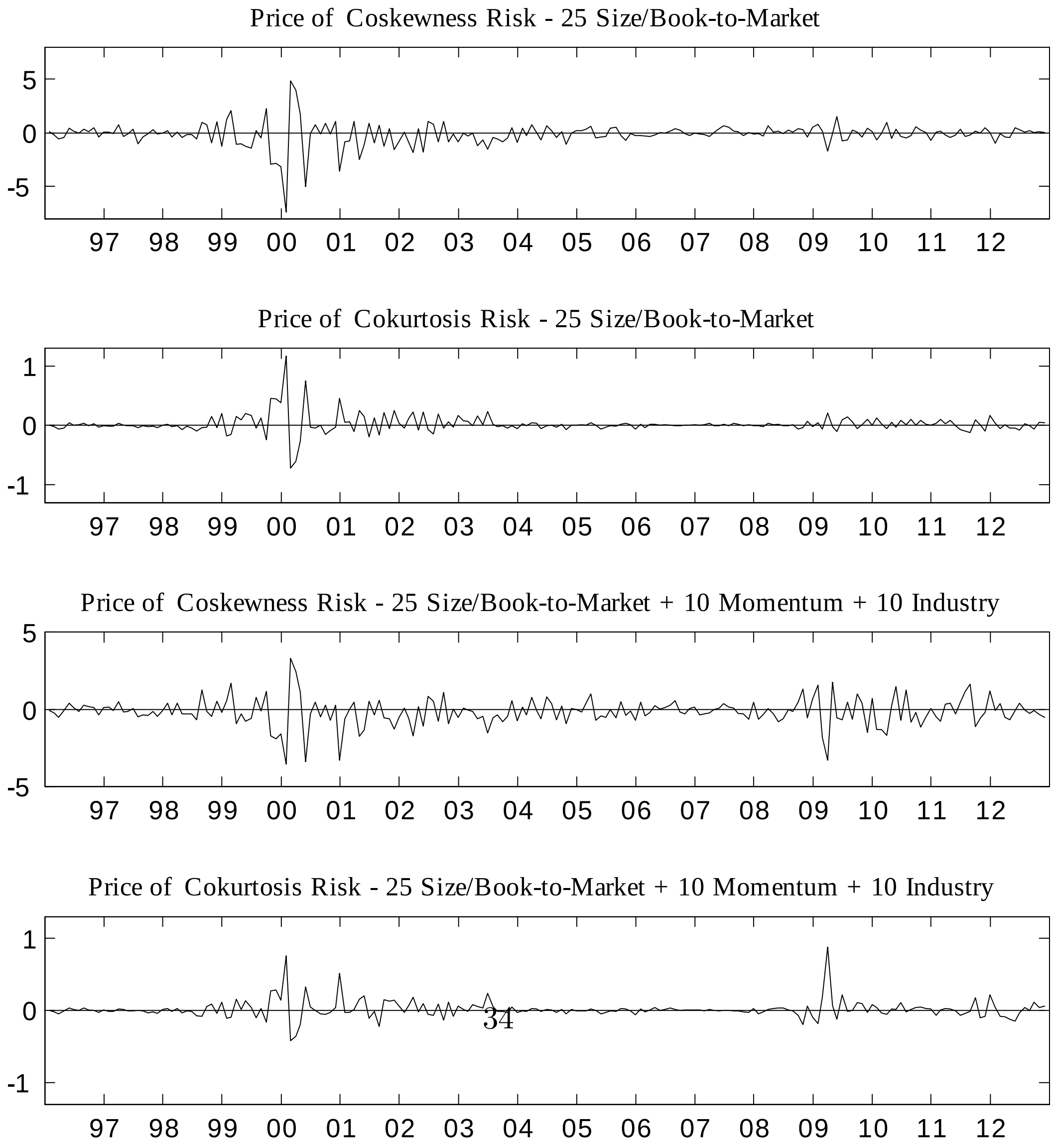

Table 1: Descriptive Statistics

We give the descriptive statistics for the physical and risk-neutral expectations and the price of co-skewness and co-kurtosis risk for monthly horizon. The mean and the standard deviation are reported based on monthly percentage returns. The second and third conditional physical moments are estimated using the autoregressive conditional volatility, skewness, and kurtosis model of Jondeau and Rockinger (2003) in which the market return follows

$$R_{m,t} = h_t z_t, \qquad z_t \sim GT(z_t|\eta_t, \lambda_t),$$

where GT denotes the generalized Student t distribution. The moment dynamics are driven by

$$\begin{aligned} h_t^2 &= a_0 + b_0^+ \left(R_{m,t-1}^+\right)^2 + b_0^- \left(R_{m,t-1}^-\right)^2 + c_0 h_{t-1}^2, \\ \eta_t &= g_\eta \left(a_1 + b_1^+ R_{m,t-1}^+ + b_1^- R_{m,t-1}^-\right), \\ \lambda_t &= g_\lambda \left(a_2 + b_2^+ R_{m,t-1}^2\right), \end{aligned}$$

where g_η and g_λ denote logistic transformations provided in Section 3.2. The risk-neutral moments are estimated using the model-free approach in Bakshi and Madan (2000) and Bakshi, Kapadia and Madan (2003). The time-varying price of risk is equal to the spread between the physical and risk-neutral moments. The sample period is from January 1996 through December 2012.

Panel A: Coskewness Risk

	$E_t^P[R_{t+1}^2]$	$E_t^Q[R_{t+1}^2]$	$E_t^P[R_{t+1}^2] - E_t^Q[R_{t+1}^2]$
mean	0.2328	0.5002	-0.2674
std	0.1680	0.4228	0.3130
skew	1.9962	3.5002	-4.4669
kurt	7.4317	20.0785	31.3785
$\rho(1)$	0.8768	0.8323	0.7339

Panel B: Cokurtosis Risk

	$E_t^P[R_{t+1}^3]$	$E_t^Q[R_{t+1}^3]$	$E_t^P[R_{t+1}^3] - E_t^Q[R_{t+1}^3]$
mean	-0.0095	-0.0228	0.0134
std	0.0146	0.0365	0.0292
skew	-3.4350	-5.4637	6.5486
kurt	16.9378	40.4634	58.4413
$\rho(1)$	0.7397	0.7632	0.6344

Table 2: Average Excess Returns and Betas

We compute the co-skewness and co-kurtosis betas for three sets of portfolios. The betas are computed for two periods: 1996-2012 and 1990-2012. The co-skewness beta, β_j^{COSK}, is computed from the regression of excess returns on market returns and squared market returns, while the co-kurtosis beta, β_j^{COKU}, is from the regression of excess returns on market returns and cubic market returns. β_j^{SKD} is defined by equation (2) and $E[R_{j,t}] - r_f$ is the average excess return. The last row of the table shows the cross-sectional correlation between the average excess returns and the betas.

		1996–2012				1990–2012			
		$E[R_{j,t}] - r_f$	β_j^{COSK}	β_j^{COKU}	β_j^{SKD}	$E[R_{j,t}] - r_f$	β_j^{COSK}	β_j^{COKU}	β_j^{SKD}
Panel A: 25 Size/Book-to-Market Portfolios									
	1. Low	0.106	-0.790	-4.173	-0.005	0.070	-1.128	-3.468	-0.006
	2	0.832	-0.441	0.930	-0.003	0.827	-0.967	0.044	-0.006
1. Small	3	0.942	-0.574	2.544	-0.006	0.866	-0.733	1.432	-0.006
	4	1.033	-0.535	3.013	-0.005	1.000	-0.909	1.744	-0.008
	5 . High	1.129	-1.571	9.517	-0.015	1.108	-1.938	8.319	-0.016
	1. Low	0.455	-0.130	-0.465	-0.001	0.485	-0.100	-2.049	-0.001
	2	0.718	-0.212	5.455	-0.002	0.670	-0.423	3.170	-0.004
2	3	0.898	-0.110	6.547	-0.001	0.903	-0.495	5.094	-0.005
	4	0.798	-1.199	12.889	-0.013	0.836	-1.419	11.039	-0.015
	5 . High	0.823	-0.712	-0.564	-0.007	0.824	-0.974	-2.625	-0.008
	1. Low	0.417	-0.624	1.684	-0.006	0.539	-0.503	0.246	-0.004
	2	0.753	0.073	6.186	0.001	0.764	-0.003	5.464	0.000
3	3	0.846	-0.066	5.464	-0.001	0.837	-0.297	4.339	-0.004
	4	0.789	0.168	7.870	0.002	0.802	0.043	6.298	0.001
	5 . High	1.058	-0.423	5.919	-0.005	1.057	-0.506	3.995	-0.005
	1. Low	0.705	-0.018	0.094	0.000	0.754	0.183	-0.414	0.002
	2	0.721	-0.454	11.396	-0.007	0.705	-0.439	9.518	-0.006
4	3	0.690	-0.786	16.924	-0.010	0.695	-0.645	13.698	-0.008
	4	0.815	-0.324	5.838	-0.004	0.803	-0.386	4.568	-0.005
	5 . High	0.636	0.032	1.158	0.000	0.663	-0.090	-0.009	-0.001
	1. Low	0.487	0.357	-4.009	0.009	0.565	0.619	-2.506	0.013
	2	0.594	-0.241	3.375	-0.005	0.644	-0.195	1.637	-0.003
5. Big	3	0.480	0.087	6.702	0.001	0.518	-0.044	6.443	-0.001
	4	0.439	-0.159	9.194	-0.002	0.475	-0.295	6.299	-0.003
	5 . High	0.420	0.363	0.630	0.004	0.572	0.148	1.295	0.001
Std			0.459	5.082	0.005		0.554	4.432	0.006
Correlation with $E[R_{j,t}] - r_f$			-0.339	0.386	-0.388		-0.324	0.354	-0.389

Table 2 – Continued

	1996–2012				1990–2012			
	$E[R_{j,t}] - r_f$	β_j^{COSK}	β_j^{COKU}	β_j^{SKD}	$E[R_{j,t}] - r_f$	β_j^{COSK}	β_j^{COKU}	β_j^{SKD}
Panel B: 10 Momentum Portfolios								
1. Low	0.063	2.506	-1.759	0.014	-0.047	1.641	-1.963	0.009
2	0.391	0.632	11.853	0.006	0.407	0.413	9.324	0.003
3	0.510	1.091	10.320	0.011	0.495	1.094	7.672	0.010
4	0.627	0.323	8.114	0.004	0.580	0.268	6.607	0.003
5	0.527	0.673	4.911	0.011	0.502	0.637	3.196	0.009
6	0.414	-0.666	10.312	-0.011	0.495	-0.552	9.252	-0.009
7	0.561	-0.146	4.071	-0.002	0.599	0.151	3.917	0.002
8	0.679	-0.313	0.261	-0.006	0.740	-0.073	0.012	-0.001
9	0.504	-0.478	-0.246	-0.008	0.586	-0.255	-0.917	-0.004
10 . High	0.870	-0.129	-13.433	-0.001	1.003	0.068	-13.228	0.001
Std		0.944	7.639	0.009		0.649	6.858	0.006
Correlation with $E[R_{j,t}] - r_f$		-0.673	-0.369	-0.420		-0.666	-0.402	-0.398
Panel C: 10 Industry Portfolios								
NoDur	0.610	-0.842	10.383	-0.011	0.649	-0.156	7.542	-0.002
Durbl	0.461	0.752	25.774	0.006	0.484	0.141	24.791	0.001
Manuf	0.693	-0.023	7.396	0.000	0.707	-0.097	6.596	-0.001
Enrgy	0.871	-0.640	6.893	-0.005	0.739	-0.664	5.196	-0.005
HiTec	0.672	1.145	-19.066	0.011	0.814	0.727	-17.159	0.006
Telcm	0.304	0.186	-7.304	0.002	0.359	-0.184	-4.641	-0.002
Shops	0.649	0.151	2.736	0.002	0.645	0.565	0.480	0.007
Hlth	0.553	-0.344	0.413	-0.004	0.671	0.489	-1.957	0.005
Utils	0.541	-0.150	1.048	-0.001	0.514	-0.187	0.558	-0.002
Other	0.380	-0.139	10.635	-0.002	0.498	-0.114	8.262	-0.002
Std		0.597	11.840	0.006		0.426	10.763	0.004
Correlation with $E[R_{j,t}] - r_f$		-0.220	-0.079	-0.157		0.293	-0.360	0.325

Table 3: Cross-Sectional Prices of Risk

We show the results of cross-sectional Fama-MacBeth regressions for monthly returns. Each month, we estimate the betas using a 60-month rolling window of monthly returns from a time series regression of the following general form

$$\begin{aligned} R_{j,t} - r_f &= \alpha_j + \beta_{j,t}^{MKT} R_{MKT,t} + \beta_{j,t}^{HML} R_{HML,t} + \beta_{j,t}^{SMB} R_{SMB,t} + \beta_{j,t}^{MOM} R_{MOM,t} \\ &+ \beta_{j,t}^{COSK} R_{MKT,t}^2 + \beta_{j,t}^{COKU} R_{MKT,t}^3 + \varepsilon_{j,t}. \end{aligned}$$

We then run the following cross-sectional regression using returns over the next month

$$\begin{aligned} R_{j,t+1} - r_f &= \lambda_{0,t} + \beta_{j,t}^{MKT} \lambda_{MKT,t} + \beta_{j,t}^{HML} \lambda_{HML,t} + \beta_{j,t}^{SMB} \lambda_{SMB,t} + \beta_{j,t}^{MOM} \lambda_{MOM,t} \\ &+ \beta_{j,t}^{COSK} \lambda_{COSK,t} + \beta_{j,t}^{COKU} \lambda_{COKU,t} + e_{j,t}. \end{aligned}$$

We report the average of the estimates and the Fama-MacBeth t-statistics with Newey-West correction for serial correlation, using 1 lag. The sample period is from January 1996 through December 2012.

Panel A: 25 Size/Book-to-Market Portfolios						
λ_0	0.837	0.785	1.070	1.173	1.265	1.294
	(1.65)	(1.38)	(2.01)	(2.41)	(2.60)	(2.58)
λ^{MKT}	-0.207	-0.331	-0.586	-0.721	-0.795	-0.812
	-(0.34)	-(0.53)	-(1.01)	-(1.39)	-(1.52)	-(1.55)
λ^{HML}				0.217	0.211	0.214
				(0.86)	(0.85)	(0.86)
λ^{SMB}				0.308	0.294	0.306
				1.150	(1.10)	(1.15)
λ^{MOM}				0.875	1.193	1.072
				(1.52)	(2.04)	(1.80)
λ^{COSK}		-0.153	-0.101		-0.099	-0.073
		-(1.98)	-(1.43)		-(1.92)	-(1.44)
λ^{COKU}			0.009			0.005
			(0.99)			(0.87)
Adj R^2	17.49	25.57	30.54	46.51	47.91	48.79
Panel B: 25 Size/Book-to-Market + 10 Momentum+10 Industry						
λ_0	0.672	0.550	0.645	0.736	0.786	0.898
	(1.69)	(1.32)	(1.57)	(2.48)	(2.60)	(2.96)
λ^{MKT}	-0.084	-0.032	-0.117	-0.235	-0.298	-0.411
	-(0.16)	-(0.06)	-(0.22)	-(0.53)	-(0.67)	-(0.94)
λ^{HML}				0.216	0.242	0.220
				(0.87)	(0.96)	(0.87)
λ^{SMB}				0.200	0.204	0.233
				(0.80)	(0.82)	(0.93)
λ^{MOM}				0.281	0.278	0.244
				(0.80)	(0.79)	(0.69)
λ^{COSK}		-0.149	-0.134		-0.086	-0.084
		-(2.46)	-(2.27)		-(2.05)	-(2.00)
λ^{COKU}			0.011			0.008
			(1.35)			(1.53)
Adj R^2	16.52	23.74	28.33	43.69	45.96	48.18

Table 4: Out-of-Sample RMSFEs

We compute the square root of mean forecasting errors (RMSFE) for different competing models: the CAPM, the model with a co-skewness premium COSK, the model with a co-kurtosis risk premium COKU, the model with market and co-skewness factors (CAPM + COSK), the model with market and co-kurtosis factors (CAPM + COKU), the model that uses the market and the alternative co-skewness measure SKD, and the Fama-French three-factor model (FF). The price of risk is estimated using cross-sectional regressions or historical risk premia. We consider three sets of portfolios: 25 size/book-to-market portfolios (Panel A), 10 momentum portfolios (Panel B), and 10 industry portfolios (Panel C). The sample period is from January 1996 through December 2012.

Panel A: 25 Size/Book-to-Market Portfolios

		Prices of Risk Estimated from Historical Risk Premia					Prices of Risk Estimated from Cross-Sectional Regressions				
		CAPM	COSK	COKU	CAPM + COSK	CAPM + COKU	CAPM	CAPM + COSK	CAPM + COKU	CAPM + SKD	FF
	1. Low	9.252	9.092	9.163	9.270	9.293	9.199	9.246	9.481	9.218	9.273
	2	7.766	7.659	7.695	7.741	7.733	7.817	7.834	7.983	7.826	7.945
1. Small	3	6.317	6.210	6.269	6.262	6.276	6.395	6.368	6.563	6.373	6.550
	4	5.908	5.756	5.837	5.786	5.829	5.988	5.933	6.054	5.943	6.194
	5 . High	6.541	6.404	6.469	6.465	6.493	6.571	6.528	6.669	6.530	6.748
	1. Low	7.914	7.749	7.768	7.905	7.882	7.911	7.963	8.071	7.945	7.984
	2	6.256	6.157	6.178	6.236	6.219	6.303	6.322	6.474	6.314	6.437
2	3	5.739	5.604	5.650	5.675	5.681	5.788	5.800	5.875	5.786	5.956
	4	5.828	5.721	5.752	5.786	5.769	5.876	5.849	5.903	5.844	6.040
	5 . High	6.661	6.521	6.527	6.593	6.565	6.695	6.671	6.814	6.672	6.864
	1. Low	7.388	7.249	7.310	7.385	7.419	7.395	7.428	7.480	7.420	7.455
	2	5.818	5.696	5.756	5.749	5.787	5.854	5.860	5.866	5.847	6.063
3	3	5.422	5.292	5.321	5.351	5.347	5.478	5.488	5.539	5.474	5.691
	4	5.509	5.381	5.428	5.439	5.472	5.509	5.546	5.538	5.524	5.754
	5 . High	5.843	5.721	5.764	5.748	5.780	5.910	5.895	5.961	5.889	6.169
	1. Low	6.614	6.564	6.589	6.618	6.644	6.686	6.729	6.684	6.716	6.806
	2	5.441	5.349	5.390	5.411	5.427	5.488	5.521	5.527	5.515	5.731
4	3	5.688	5.632	5.631	5.708	5.681	5.692	5.682	5.688	5.675	5.970
	4	5.297	5.185	5.291	5.215	5.315	5.339	5.363	5.360	5.333	5.573
	5 . High	5.879	5.876	5.867	5.926	5.919	5.884	5.872	5.842	5.865	6.102
	1. Low	4.910	4.960	4.900	4.970	4.954	5.027	5.057	5.063	5.064	5.082
	2	4.556	4.492	4.496	4.530	4.516	4.655	4.676	4.587	4.659	4.826
5. Big	3	4.923	4.845	4.829	4.888	4.869	4.961	4.980	4.952	4.969	5.148
	4	4.845	4.956	4.921	4.973	4.933	4.897	4.959	4.906	4.941	5.075
	5 . High	5.770	5.798	5.802	5.862	5.842	5.819	5.952	5.928	5.881	5.926
	Average	6.083	5.995	6.024	6.060	6.066	6.125	6.141	6.192	6.129	6.294

Table 4 – Continued

Panel B: 10 Momentum Portfolios

	Prices of Risk Estimated from Historical Risk Premia					Prices of Risk Estimated from Cross-Sectional Regressions				
	CAPM	COSK	COKU	CAPM + COSK	CAPM + COKU	CAPM	CAPM + COSK	CAPM + COKU	CAPM + SKD	FF
1. Low	10.774	10.810	10.712	11.009	10.927	10.708	10.751	10.772	10.760	11.019
2	7.607	7.642	7.679	7.719	7.770	7.563	7.550	7.663	7.586	7.859
3	6.311	6.378	6.469	6.436	6.539	6.323	6.257	6.302	6.324	6.537
4	5.388	5.379	5.379	5.420	5.431	5.385	5.342	5.358	5.405	5.684
5	4.844	4.835	4.765	4.918	4.816	4.856	4.824	4.810	4.930	5.051
6	4.641	4.623	4.578	4.689	4.620	4.645	4.608	4.681	4.649	4.818
7	4.426	4.374	4.309	4.441	4.355	4.460	4.385	4.390	4.421	4.664
8	4.386	4.384	4.370	4.389	4.358	4.444	4.346	4.355	4.375	4.658
9	4.795	4.728	4.772	4.804	4.825	4.779	4.720	4.780	4.710	5.038
10 . High	6.796	6.713	6.995	6.732	6.989	6.875	6.794	6.927	6.810	7.021
Average	5.997	5.987	6.003	6.056	6.063	6.004	5.958	6.004	5.997	6.235

Panel C: 10 Industry Portfolios

	Prices of Risk Estimated from Historical Risk Premia					Prices of Risk Estimated from Cross-Sectional Regressions				
	CAPM	COSK	COKU	CAPM + COSK	CAPM + COKU	CAPM	CAPM + COSK	CAPM + COKU	CAPM + SKD	FF
NoDur	3.902	3.942	3.939	3.973	3.958	3.896	3.893	3.924	3.867	3.891
Durbl	7.895	7.818	7.848	7.910	7.912	7.842	7.830	7.844	7.857	7.872
Manuf	5.354	5.370	5.398	5.408	5.427	5.338	5.353	5.367	5.339	5.328
Enrgy	5.973	5.866	6.310	5.870	6.293	5.971	5.903	5.969	5.916	5.964
HiTec	8.213	8.260	8.205	8.284	8.294	8.200	8.300	8.308	8.290	8.152
Telcm	5.858	5.786	5.941	5.883	6.018	5.828	5.897	5.916	5.899	5.786
Shops	4.833	4.790	4.772	4.811	4.801	4.833	4.854	4.851	4.851	4.816
Hlth	4.319	4.393	4.367	4.377	4.346	4.307	4.318	4.354	4.298	4.314
Utils	4.348	4.391	4.493	4.402	4.501	4.342	4.338	4.317	4.302	4.333
Other	5.679	5.647	5.641	5.728	5.693	5.630	5.652	5.621	5.690	5.633
Average	5.637	5.626	5.691	5.664	5.724	5.618	5.634	5.647	5.631	5.609

Table 5: Out of Sample $R-$square (1996–2009)

We test the out of sample performance of different models using out of sample $R-$squares. We compute out of sample $R-$squares for each portfolio and report the average. We consider three sets of portfolios and conduct a subsample analysis. The first column shows the average of out of sample $R-$squares using the entire sample period, whereas the second and third column are based on the first and second half of the sample, respectively. In the fourth (fifth) column, we retain only years in which the cumulative stock market return over the year is negative (positive).

	Full Sample	1996 to 2004	2005 to 2012	Negative ret years	Positive ret years
Panel A: 10 Industry Portfolios					
Prices of Risk Estimated from Historical Risk Premia					
CAPM	0.597	0.899	-0.107	2.557	-0.424
COSK	0.750	1.110	0.259	2.181	-0.591
COKU	-1.561	0.809	-3.778	-2.068	-1.765
CAPM + COSK	-0.517	0.258	-1.713	-0.071	-0.839
CAPM + COKU	-2.600	0.519	-5.852	-4.291	-1.661
Prices of Risk Estimated from Cross-Sectional Regressions					
CAPM	1.219	0.922	1.181	6.358	-1.711
CAPM + COSK	0.748	-0.426	1.705	3.808	-0.718
CAPM + COKU	0.253	-0.113	0.815	3.811	-1.316
CAPM +SKD	0.973	-0.345	2.323	3.014	-0.088
FF	1.549	1.360	1.299	5.595	-0.922
Panel B: 25 Size/Book-to-Market Portfolios					
Prices of Risk Estimated from Historical Risk Premia					
CAPM	-0.402	-0.836	-0.152	1.886	-1.678
COSK	2.393	1.661	3.212	4.401	1.242
COKU	1.496	-0.109	3.060	4.812	-0.399
CAPM + COSK	0.376	0.440	0.301	2.968	-0.995
CAPM + COKU	0.189	0.073	0.224	3.615	-1.619
Prices of Risk Estimated from Cross-Sectional Regressions					
CAPM	-1.944	-6.027	2.116	4.494	-6.068
CAPM + COSK	-2.472	-5.550	0.541	4.816	-7.141
CAPM + COKU	-3.931	-6.881	-1.051	3.765	-9.021
CAPM +SKD	-2.061	-5.529	1.428	4.647	-6.387
FF	-7.967	-17.323	0.808	-5.007	-10.520

Table 5 – Continued

	Full Sample	1996 to 2004	2005 to 2012	Negative ret years	Positive ret years
Panel C: 10 Momentum Portfolios					
Prices of Risk Estimated from Historical Risk Premia					
CAPM	-0.048	0.331	-0.381	0.844	-0.340
COSK	0.482	1.849	-0.556	3.754	-0.880
COKU	0.101	2.824	-2.709	1.587	0.045
CAPM + COSK	-1.765	0.197	-3.463	-0.609	-2.100
CAPM + COKU	-1.717	1.708	-5.223	-2.311	-0.742
Prices of Risk Estimated from Cross-Sectional Regressions					
CAPM	-0.490	-1.906	1.066	6.546	-3.902
CAPM + COSK	1.386	0.618	2.328	6.360	-1.051
CAPM + COKU	-0.110	-1.800	1.618	4.824	-2.555
CAPM +SKD	-0.090	-0.179	0.258	6.821	-3.494
FF	-8.788	-9.242	-8.503	10.226	-18.611

Table 6: Out of Sample $R-$squares Using Alternative Measures of Physical Moments

We test the out of sample performance of different models using out of sample $R-$squares. The prices of risk are estimated from historical risk premia and the one-step ahead forecast of the physical second and third moments– needed to compute the price of co-skewness and co-kurtosis risk– are estimated from the following monthly regressions

$$\kappa_t^2 = a_0 + a_1 \kappa_{t-1}^2 + u_t^\sigma,$$

and

$$\kappa_t^3 = b_0 + b_1 \kappa_{t-1}^3 + u_t^\sigma,$$

where $\kappa_t^2 = \sum_{d \in t} R_{d,t}^2$, $\kappa_t^3 = \sum_{d \in t} R_{d,t}^3$, and $R_{d,t}$ is the daily return in day d of month t.
We compute out of sample $R-$squares for each portfolio and report the average. We consider three sets of portfolios. The first column shows the average of out of sample $R-$squares using the entire sample period, whereas the second and third column are based on the first and second half of the sample, respectively. In the fourth (fifth) column, we retain only years in which the cumulative stock market return over the year is negative (positive).

	Full Sample	1996 to 2004	2005 to 2012	Negative ret years	Positive ret years
Panel A: 10 Industry Portfolios					
CAPM	0.597	0.899	-0.107	2.557	-0.424
COSK	0.941	1.020	0.932	2.599	-0.608
COKU	-0.512	-0.183	-0.861	0.614	-1.795
CAPM + COSK	-0.349	0.143	-1.073	0.344	-0.889
CAPM + COKU	-1.199	-0.012	-2.859	-1.556	-1.187
Panel B: 25 Size/Book-to-Market Portfolios					
CAPM	-0.402	-0.836	-0.152	1.886	-1.678
COSK	1.998	1.504	2.557	4.003	0.822
COKU	1.298	-0.859	3.496	4.948	-0.842
CAPM + COSK	-0.068	0.221	-0.387	2.542	-1.477
CAPM + COKU	0.018	-0.969	1.029	3.944	-2.112
Panel C: 10 Momentum					
CAPM	-0.048	0.331	-0.381	0.844	-0.340
COSK	0.217	1.708	-0.935	4.511	-1.826
COKU	0.665	1.069	0.483	3.840	-0.544
CAPM + COSK	-2.062	0.028	-3.869	0.165	-3.099
CAPM + COKU	-1.000	0.083	-1.844	0.039	-1.159

Table 7: Out of Sample $R-$square Using the VIX and the SKEW Index

We test the out of sample performance of different models using out of sample $R-$squares. The risk-neutral second and third moments are proxied by the VIX and the SKEW indexes, respectively. The second and third conditional physical moments are estimated using the autoregressive conditional volatility, skewness, and kurtosis model of Jondeau and Rockinger (2003). We compute out of sample $R-$squares for each portfolio and report the average. We consider three sets of portfolios. The first column shows the average of out of sample $R-$squares using the entire sample period, whereas the second and third column are based on the first and second half of the sample, respectively. In the fourth (fifth) column, we retain only years in which the cumulative stock market return over the year is negative (positive).

	Full Sample	1990 to 2001	2002 to 2012	Negative ret years	Positive ret years
Panel A: 10 Industry Portfolios					
Prices of Risk Estimated from Historical Risk Premia					
CAPM	0.665	1.100	0.174	2.177	-0.206
COSK	0.828	-0.027	1.844	3.828	-1.453
COKU	-7.465	-8.266	-5.215	-3.153	-10.928
CAPM + COSK	0.367	0.842	-0.241	1.406	-0.275
CAPM + COKU	-6.994	-5.874	-7.150	-5.580	-8.346
Prices of Risk Estimated from Cross-Sectional Regressions					
CAPM	0.273	-0.686	1.343	4.759	-2.611
CAPM + COSK	-0.806	-2.815	1.462	1.808	-2.319
CAPM + COKU	-0.642	-2.116	1.086	2.671	-2.427
CAPM +SKD	-0.278	-2.225	1.902	1.438	-1.397
FF	0.591	-0.098	1.405	4.515	-2.040
Panel B: 25 Size/Book-to-Market					
Prices of Risk Estimated from Historical Risk Premia					
CAPM	-0.257	0.339	-0.878	1.508	-1.289
COSK	1.434	-0.121	2.683	4.414	-0.464
COKU	-0.585	-2.857	1.059	6.698	-5.111
CAPM + COSK	0.391	0.919	-0.140	2.196	-0.628
CAPM + COKU	-1.049	-1.184	-1.174	4.503	-4.289
Prices of Risk Estimated from Cross-Sectional Regressions					
CAPM	-2.982	-8.877	1.873	4.443	-7.946
CAPM + COSK	-3.339	-8.377	0.765	4.775	-8.757
CAPM + COKU	-4.155	-6.920	-1.895	3.491	-9.354
CAPM +SKD	-2.891	-8.426	1.676	4.632	-7.932
FF	-7.573	-13.629	-3.218	-2.881	-11.095

Table 7 − Continued

	Full Sample	1990 to 2001	2002 to 2012	Negative ret years	Positive ret years
Panel C: 10 Momentum					
Prices of Risk Estimated from Historical Risk Premia					
CAPM	-0.051	0.091	-0.167	0.678	-0.302
COSK	0.573	-0.412	1.708	4.962	-1.639
COKU	-4.366	-4.432	-4.551	1.745	-7.328
CAPM + COSK	-0.806	-0.169	-1.182	0.547	-1.377
CAPM + COKU	-5.104	-3.604	-6.688	-2.404	-6.266
Prices of Risk Estimated from Cross-Sectional Regressions					
CAPM	-1.659	-5.729	2.125	5.424	-5.524
CAPM + COSK	-12.285	-29.002	2.499	-2.169	-18.103
CAPM + COKU	-24.962	-56.206	0.981	-13.084	-32.734
CAPM +SKD	-10.924	-24.375	1.132	-0.305	-17.008
FF	-13.222	-13.828	-12.750	8.128	-25.633