# Option-Based Estimation of the Price of Co-Skewness and Co-Kurtosis Risk

## Peter Christoffersen, Mathieu Fournier, Kris Jacobs and Mehdi Karoui

## CREATES Research Paper 2015-54

# Option-Based Estimation of the Price of Co-Skewness and Co-Kurtosis Risk*

Peter Christoffersen

University of Toronto,

CBS and CREATES

Mathieu Fournier

HEC Montréal


Kris Jacobs

University of Houston

Mehdi Karoui

OMERS

September 1, 2015

## Abstract

We show that the prices of risk for 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. The 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. Models with 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

# Contents

# 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 out-of-sample 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 mixed. For instance, the existing literature is divided on the performance of aggregate consumption growth, the most important state variable suggested by theory.[1] 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 in-sample, but the lack of economic and theoretical foundations is cause for concern.[2]

Given 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.[3] 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

---

[1] Early studies such as Breeden, Gibbons, and Litzenberger (1989) and Mankiw and Shapiro (1986) argue that the consumption-based model performs poorly in pricing the cross-section of stock returns. For a more positive assessment of the performance of the conditional consumption-based model, see for instance Bansal, Dittmar and Lundblad (2005), Dittmar and Lundblad (2015), and Lettau and Ludvigson (2001).

[2] 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.

[3] In a related literature, Ang, Hodrick, Xing, and Zhang (2006) analyze the performance of volatility as a pricing factor. Schneider, Wagner, and Zechner (2015) offers an insightful comparison of the role of volatility and skewness in cross-sectional studies.

(2002), Guidolin and Timmermann (2008), and Scott and Horvath (1980)).[4] 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.

One possible drawback of co-skewness and co-kurtosis as cross-sectional pricing factors is measurement. Measurement is especially difficult when analyzing *conditional* co-skewness and co-kurtosis.[5] Most existing papers estimate and test the importance of co-skewness and co-kurtosis using two-stage cross-sectional regressions. For a classical example of this type of conditional analysis, see for instance 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 the estimation of betas in the first-stage regression is noisy, and these errors carry over in the second-stage cross-sectional regression.[6] 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 reason 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.[7] SDFs that are higher-order functions of the market return lead to progressively

---

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

[5]Kraus and Litzenberger (1976) provide an unconditional empirical analysis of co-skewness.

[6]See, e.g. Jagannathan and Wang (1998), Shanken (1992), Kan and Zhang (1999), Kleibergen (2009), and Gospodinov, Kan, and Robotti (2014).

[7]See Dittmar (2002) for an investigation of higher moments in cross-sectional pricing using this approach.

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. Remarkably, we can show that the price of 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 implies that the price of risk corresponds to 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. While information from index options is not particularly useful in the linear SDF case, we show that whenever the SDF is nonlinear then information from index option prices can be used to pin down the price of risk of the nonlinear factor. Fortunately, we have particularly rich option information on the market index, as index options are among the most heavily traded contracts on the market. This makes our theoretical results very practical to implement.

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 almost 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 larger in absolute value than the traditional estimates obtained using a two-stage Fama-MacBeth approach. While the average prices of risk obtained using the Fama-MacBeth approach have the theoretically anticipated signs, the estimates have the opposite sign in many months, which explains the smaller averages. 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-French 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. In Section 3 we motivate our approach by discussing traditional regression-based estimates. Section 4 presents an empirical investigation of co-

---

See Bakshi, Madan, and Panayotov (2010) for evidence that the pricing kernel is U-shaped as a function of market return.

skewness risk using our approach. Section 5 investigates co-kurtosis risk. Section 6 concludes.

# 2  Measuring Market Risks:  An Option-Based Approach

In this section we provide an overview of multifactor asset pricing 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

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

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) among others.

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 provided that information on option prices is available.

Denote the return on a stock by $R_j$ and the return on the market index by $R_m$. The existing literature contains several measures of co-skewness risk, which all capture covariation

6

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 analysis but consider four different co-skewness measures in their empirical analysis. Our measure of co-skewness risk is the 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** *In the absence of arbitrage opportunities, if the stochastic discount factor (SDF) has the following form*

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

*then the cross-sectional pricing restriction on stock returns is*

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

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

$$\lambda_t^{MKT} = E_t^P(R_{m,t+1}) - R_{f,t},\tag{4}$$

*and the price of co-skewness risk, $\lambda_t^{COSK}$, is*

$$\lambda_t^{COSK} = E_t^P(R_{m,t+1}^2) - E_t^Q(R_{m,t+1}^2).\tag{5}$$

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

**Proof.** 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 mean zero risk factors $\mathbf{f}$ satisfying

$$E_t^P(R_{j,t+1}) - R_{f,t} = \boldsymbol{\lambda}_t'\boldsymbol{\beta}_{j,t},\tag{6}$$

where $\boldsymbol{\lambda}_t' = \frac{-1}{a_t}\mathbf{b}_t'E_t^P(\mathbf{f}_{t+1}\mathbf{f}_{t+1}')$, $(1+R_{f,t}) = \frac{1}{a_t} = \frac{1}{E_t^P(m_{t+1})}$, and $\boldsymbol{\beta}_{j,t} = \left[E_t^P(\mathbf{f}_{t+1}\mathbf{f}_{t+1}')\right]^{-1}E_t^P(\mathbf{f}_{t+1}R_{j,t+1})$, see for instance Cochrane (2005). Since the pricing kernel prices all the assets, the above

equation also holds for any contingent claim with payoff $\Psi$, which can be a function of the market index return or of the stock return. Consequently, applying equation (6) to $\Psi$ gives

$$E_t^P \left( R_{\Psi,t+1} \right) - R_{f,t} = E_t^P \left( \frac{\Psi_{t+1} - P_t}{P_t} \right) - R_{f,t} = \boldsymbol{\lambda}_t' \boldsymbol{\beta}_{\Psi,t}, \tag{7}$$

where $P_t$ is the price of the contingent claim $\Psi$ and $R_\Psi$ is the return on the contingent claim. Using the definition of $\boldsymbol{\beta}_{\Psi,t}$ we have

$$E_t^P \left( \frac{\Psi_{t+1} - P_t}{P_t} \right) - R_{f,t} = \boldsymbol{\lambda}_t' \left[ E_t^P (\mathbf{f}_{t+1} \mathbf{f}_{t+1}') \right]^{-1} E_t^P \left( \mathbf{f}_{t+1} \left( \frac{\Psi_{t+1} - P_t}{P_t} \right) \right). \tag{8}$$

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

$$
\begin{aligned}
E_t^P(\Psi_{t+1}) - P_t \left( 1 + R_{f,t} \right) &= \boldsymbol{\lambda}_t' \left[ E_t^P (\mathbf{f}_{t+1} \mathbf{f}_{t+1}') \right]^{-1} E_t^P \left( \mathbf{f}_{t+1} \Psi_{t+1} \right) \\
&= \boldsymbol{\lambda}_t' \tilde{\boldsymbol{\beta}}_{\Psi,t},
\end{aligned} \tag{9}
$$

where $\tilde{\boldsymbol{\beta}}_{\Psi,t}$ is from the projection of $\Psi$ on $\mathbf{f}$. The no-arbitrage condition ensures the existence of at least one risk-neutral measure $Q$ such that $P_t = \frac{E_t^Q(\Psi_{t+1})}{(1+R_{f,t})}$. Therefore, we get

$$E_t^P \left( \Psi_{t+1} \right) - E_t^Q \left( \Psi_{t+1} \right) = \boldsymbol{\lambda}_t' \tilde{\boldsymbol{\beta}}_{\Psi,t}. \tag{10}$$

To obtain the result in equation (4) from equation (10), we now consider the contingent claim $\Psi_{t+1} \equiv R_{m,t+1}$. If a return is also a factor, it has a loading of one onto itself and zero onto the other factors. Given the SDF (2), this gives $\tilde{\boldsymbol{\beta}}_{\Psi,t} = [1 \ 0]'$ and equation (10) reduces to equation (4). Similarly, using $\Psi_{t+1} \equiv R_{m,t+1}^2$, we obtain $\tilde{\boldsymbol{\beta}}_{\Psi,t} = [0 \ 1]'$ which applied to equation (10) gives the result in equation (5). ∎

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. 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.

The spread between the physical and risk neutral market variance is often termed the variance risk premium and it has been found to be one of the best predictor of market returns. See, for example, Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Marrone, Xu and

Zhou (2014). Our analysis above suggests that the variance risk premium is a predictor of market returns because it provides information about the price of co-skewness risk.

Proposition 1 allows for separate identification of the price of covariance $(\lambda_t^{MKT})$ and co-skewness $(\lambda_t^{COSK})$ 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. In our empirical implementation, we show that an additional advantage of our approach is that the period-by period estimates of the conditional price of risk are rather reliable and precise, in contrast with the estimates obtained using the regression-based approach.

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.

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 as in Bakshi and Madan (2000) and Bakshi, Kapadia, and Madan (2003).

While our approach to estimating the price of co-skewness risk is different from the existing literature and the betas are defined (and/or scaled) differently, the implications for the risk premia on the assets are of course the same. Using the fact that $E_t^P(R_{m,t+1}) - R_{f,t} = \lambda_t^{MKT}$ 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

$$E_t^P(R_{j,t+1}) - R_{f,t} = \beta_{j,t}^{MKT}\left[E_t^P(R_{m,t+1}) - R_{f,t}\right] + \beta_{j,t}^{COSK}\left[E_t^P(R_{m,t+1}^2) - E_t^Q(R_{m,t+1}^2)\right], \quad (11)$$

9

which can also be written as

$$E_t^P(R_{j,t+1}) - R_{f,t} = c_t + \beta_{j,t}^{MKT} E_t^P(R_{m,t+1}) + \beta_{j,t}^{COSK} E_t^P\left(R_{m,t+1}^2\right), \tag{12}$$

where $c_t = -\beta_{j,t}^{MKT} R_{f,t} - \beta_{j,t}^{COSK} E_t^Q(R_{m,t+1}^2)$. Equation (12) 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 no-arbitrage restrictions on contingent claims prices so that the pricing kernel prices all assets in the economy. These additional innocuous restrictions lead to a very simple estimator of the price of risk.

## 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

$$\begin{aligned}
m_{t+1} &= a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right) \\
&\quad + b_{3,t}\left(R_{m,t+1}^3 - E_t^P(R_{m,t+1}^3)\right).
\end{aligned} \tag{13}$$

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 by $\beta_{j,t}^{COKU}$.

10

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

**Proposition 2** *In the absence of arbitrage opportunities, if the stochastic discount factor (SDF) has the following form:*

$$m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P(R_{m,t+1}^2)\right)$$
$$+ b_{3,t}\left(R_{m,t+1}^3 - E_t^P(R_{m,t+1}^3)\right), \tag{14}$$

*then the cross-sectional restriction on stock returns is*

$$E_t^P(R_{j,t+1}) - R_{f,t} = \lambda_t^{MKT}\beta_{j,t}^{MKT} + \lambda_t^{COSK}\beta_{j,t}^{COSK} + \lambda_t^{COKU}\beta_{j,t}^{COKU}, \tag{15}$$

*where $\beta_{j,t}^{MKT}$, $\beta_{j,t}^{COSK}$, and $\beta_{j,t}^{COKU}$ are from the projection of asset returns on $R_{m,t+1}$, $R_{m,t+1}^2$ and $R_{m,t+1}^3$, respectively. The prices of covariance, $\lambda_t^{MKT}$, and co-skewness risk $\lambda_t^{COSK}$ are*

$$\lambda_t^{MKT} = E_t^P(R_{m,t+1}) - R_{f,t}, \tag{16}$$

$$\lambda_t^{COSK} = E_t^P(R_{m,t+1}^2) - E_t^Q(R_{m,t+1}^2), \tag{17}$$

*and the price of co-kurtosis risk, $\lambda_t^{COKU}$, is*

$$\lambda_t^{COKU} = E_t^P(R_{m,t+1}^3) - E_t^Q(R_{m,t+1}^3), \tag{18}$$

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

**Proof.** The structure of the proof largely follows the proof of Proposition 1. Given equation (14), applying equation (10) for $\Psi_{t+1} \equiv R_{m,t+1}$ as in Proposition 1, we recover equation (16), and applying equation (10) for $\Psi_{t+1} \equiv R_{m,t+1}^2$, we recover equation (17). In addition, applying equation (10) for $\Psi_{t+1} \equiv R_{m,t+1}^3$, we obtain equation (18). This again uses the results that a return which is also a factor has a loading of one onto itself and zero on the other factors. ∎

Proposition 2 shows that the price of co-kurtosis risk is equal to the spread between the market physical and risk-neutral third moments. Existing empirical 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

11

positive price of co-kurtosis risk. This is 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 higher than the third order, and therefore we do not extend our empirical analysis beyond the cubic SDF. While the empirical focus of this paper is on co-skewness and co-kurtosis risk, our approach can be used for virtually any source of risk that is an integrable function of the market return. This does not just include expectations of 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 case which nests the results in Propositions 1 and 2 of the previous Section.

**Proposition 3** *In the absence of arbitrage opportunities, if the stochastic discount factor (SDF) has the following form:*

$$m_{t+1} = a_t + \sum_k b_{k,t} \left( G_k(R_{m,t+1}) - E_t^P \left( G_k(R_{m,t+1}) \right) \right) + \sum_l c_{l,t} \left( f_{l,t+1} - E_t^P(f_{l,t+1}) \right), \quad (19)$$

*then the cross-sectional pricing restriction for stock returns is*

$$E_t^P(R_{j,t+1}) - R_{f,t} = \sum_k \lambda_t^k \beta_{j,t}^k + \sum_l \gamma_t^l \beta_{j,t}^l, \quad (20)$$

*where $\beta_{j,t}^k$ and $\beta_{j,t}^l$ are from the projection of asset returns on $G_k(R_{m,t+1})$ and $f_{l,t+1}$ respectively, and $\gamma^l$ 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, $\lambda_t^k$, is*

$$\lambda_t^k = E_t^P(G_k(R_{m,t+1})) - E_t^Q(G_k(R_{m,t+1})), \quad (21)$$

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

    **Proof.** The structure of the proof is again similar to the proof of Proposition 1. Given equation (19), then applying equation (10) for $\Psi_{t+1} \equiv G_k(R_{m,t+1})$, we obtain equation (21). ∎

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. The proposition also demonstrates that we can easily allow for factors that are not necessarily functions of the market return.

# 3    Regression-Based Estimates of the Price of Co-Skewness Risk

The main purpose of the paper is to present estimates of the price of co-skewness and co-kurtosis risk using the newly proposed methods in Section 2. Before we do so, we first present traditional regression-based estimates of the price of co-skewness risk. The purpose of this exercise is twofold. First, these estimates serve as a benchmark for our new estimates. Second, we will use these results to highlight problems that arise with estimating the price of risk for higher moments. Here we limit ourselves to a discussion of regression-based estimates of co-skewness risk. We briefly discuss regression-based estimates of co-kurtosis risk in Section 5.2 below.

## 3.1    Portfolios Sorted on Co-Skewness Exposure

We first present estimates of co-skewness risk obtained using portfolios that are formed by sorting the cross-section of stocks according to co-skewness exposure. We present results for the period 1986-2012, which coincides with the sample period we will use for our option-based estimates. It could be argued that this sample period is relatively short to reliably estimate the price of co-skewness risk using a regression approach. We therefore also report on the period 1966-2012. We present results from Fama-MacBeth regressions using the classical setup. We obtain betas using sixty monthly returns, and subsequently we run a cross-sectional regression for the next month.

For this exercise we consider the entire cross-section of CRSP stocks and restrict our sample to common shares. In any given month, we do not use firms with fewer than seventeen consecutive monthly observations preceding the estimation month or with fewer than thirty-six observations in total (i.e. sixty percent of the length of the estimation window). These filters ensure that the firms in our sample have sufficiently rich data available for reliably estimating the exposures. Based on the co-skewness and CAPM betas, we double sort stocks into quintile portfolios and compute the value-weighted monthly returns on the resulting

twenty-five portfolios. We use these portfolios as test assets to estimate the price of co-skewness risk and study patterns in co-skewness exposure.

Table 1 presents the results. The two left-side columns of Panel A report on the 1986-2012 sample. The two right-side columns report on the 1966-2012 sample. For both sample periods, we report on two models. The first model incorporates co-skewness exposure and exposure to the market factor. The second model also includes the Fama-French and momentum factors.

The estimates of the price of co-skewness risk are negative in all four cases, consistent with theory, and they are of a similar order of magnitude. However, none of the estimates are statistically significant. These results are not surprising. It is well known that using a regression-based approach, it is difficult to obtain precise estimates of the price of co-skewness risk, and the price of higher moment risk more generally.

Panel B provides some insight into the underlying reason for these findings. We report on the co-skewness and CAPM pre-formation and post-formation exposures for both sample periods. The first four columns from the left report on co-skewness exposures. The four columns on the right present the exposure to covariance risk. Note that by definition the pattern for the pre-formation exposure is monotonic for both factors. The spread in the exposures is also substantial.

More interestingly, the full sample post-formation results suggest an important difference between market exposure and co-skewness exposure. In the case of market exposure, the differences between the exposures of the quintile portfolios are of course smaller than in the pre-formation case, but the post-formation differences are substantial and the exposures are monotonic as a function of the portfolios. For the co-skewness exposure, this is not the case. The reason is that to reliably estimate higher moment risk exposure, it is important to observe extreme market and stocks returns. In practice, the relative infrequent occurrence of extreme returns renders the estimation of co-skewness betas difficult, and this explains the post-formation patterns.

A potential solution is to use longer estimation windows to increase the occurrence of extreme observations. However, it is well-known that co-skewness exposure is not very persistent, and in this case the longer regression samples do not lead to better estimates of co-skewness exposure. In summary, there are inherent problems with the use of regression-based methods to estimate higher moment risk.

## 3.2 Other Portfolios

One possible reasons for the findings in Table 1 is that it is difficult to measure the co-skewness exposure for individual stocks. We therefore also present regression-based results obtained using portfolios that are not formed and sorted based on co-skewness exposure. This means that we do not have to compute the co-skewness exposure of individual stocks. Instead we directly compute the exposure of the portfolios, and these estimates are less noisy.

We again present results for 1986-2012 and 1966-2012 using Fama-MacBeth regressions. We obtain betas using sixty monthly returns, and subsequently we run a cross-sectional regression for the next month. Similar to Table 1, Table 2 reports results for two models. The first model incorporates co-skewness exposure and exposure to the market factor. The second model also includes the Fama-French (1993) size and book-to-market factors, and the momentum factor. Figure 1 reports on a third model, the univariate model that exclusively contains co-skewness exposure.

For each regression, following Fama and MacBeth (1973), we report the average of the cross-sectional regression estimates as well as the t-statistics on these averages. We report on four cross-sectional datasets that are commonly used in the existing literature. We use portfolios formed on size and book-to-market ratio, on size and momentum, on size and short-term reversal, and portfolios formed on size and long-term reversal. 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.

Consider first the results for 1986-2012 in Panel A of Table 2 and Figure 1. An important conclusion is that the estimates of the price of co-skewness risk critically depend on the assets used in estimation. For the univariate model displayed in Figure 1 (note that the exposures for this model are estimated using bivariate regressions with the market factor and the market factor squared), the estimate of the price of co-skewness risk is $-0.084$ when using the twenty-five size and book-to-market portfolios.[8] When using the twenty-five size and momentum portfolios, the estimate is $-0.182$. However, when using the size and short-term reversal portfolios and the size and long-term reversal portfolios, the estimates are positive. The only estimate that is statistically significant is the one obtained using the twenty-five size and momentum portfolios. Panel A of Table 2 indicates that when including the market factor in the regressions, the results do not change much. The estimates for the size and short-term reversal portfolios and the size and long-term reversal portfolios are now negative

---

[8]For expositional convenience we report the estimated prices of co-skewness and co-kurtosis risk, as well as the moments used to construct these estimates, in percentage terms, i.e. multiplied by 100.

but they are not statistically significant.

Including additional factors in the cross-sectional model does not change this conclusion either. Table 2 reports results for the price of co-skewness risk when the Fama-French factors as well as the momentum factor are included in the regressions. The resulting estimates are smaller in absolute value and are never statistically significant.

As mentioned before, we focus on the 1986-2012 period to compare the results to our newly proposed estimates, which are limited to this sample period because of the availability of risk-neutral second moments. Panel B of Table 2 also reports results for the longer 1966-2012 period. The resulting estimates of the price of co-skewness risk are very similar to those obtained for the 1986-2012 period, and also strongly differ across test assets.

Our first conclusion is that the choice of test assets is important for the point estimates and significance of the price of co-skewness risk.

Figure 2 reports on the model that includes the co-skewness and market factors for the 1986-2012 sample. We report the time-series of the month-by-month cross-sectional regression estimates of the price of co-skewness risk. The estimates for the price of co-skewness risk reported in the first model in Panel A of Table 2 are the averages of the time series in Figure 1. Figure 2 yields several important conclusions. Based on the results in Figure 1 and Table 2, we concluded that the choice of test assets substantially impacts the estimates of the price of risk. Figure 2 instead suggests substantial commonality between test assets in the month-by-month estimates of the price of risk. The four time series in Figure 2 clearly have common features. Table 2 indicates that the only test assets that yield a significantly negative price of co-skewness risk are the twenty-five size and momentum portfolios. Figure 2 indicates that this can be explained by the fact that the regression estimates are noisy, and the estimates for these test assets vary less over time compared to the estimates for other test assets, even though the monthly estimates are also often positive.

The essence of the Fama-MacBeth procedure is of course to estimate the price of risk by averaging the time series of cross-sectional estimates. The fact that the estimates in Figure 2 are positive for some months therefore does not constitute a problem in itself. But it is clear that the cross-sectional estimates vary a lot over time, and that they are often positive, even when the averages reported in Panel A of Table 2 are negative. Figure 2 therefore suggests that noise in the month-by-month estimates is an important problem with regression-based estimation of the price of co-skewness risk.

# 4  New Estimates of the Price of Co-Skewness Risk

We now present estimates of the price of co-skewness risk using the estimators presented in Proposition 1. 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, $E_t^P(R_{m,t+1}^2)$. We first discuss the estimation of these moments. Subsequently we estimate the price of co-skewness risk and discuss the differences between our new estimates and conventional regression-based estimates.

## 4.1  Estimating the Risk-Neutral Second Moment

We estimate the risk-neutral variance in two ways. In our benchmark analysis, we use the square of the VIX index as our estimate for the risk-neutral variance.[9] The VIX provides a very simple benchmark because the data are readily available from the Chicago Board of Options Exchange (CBOE). Using the VIX has a number of advantages. The construction of the VIX is exogenous to our experiment, and so it is not possible to design it to maximize performance. Even more importantly, the VIX is available for a longer sample period than the available alternatives. We use data for the ticker VXO throughout and obtain data for the period January 1986 to December 2012. For existing studies that use the VIX squared as a proxy for the expected risk-neutral second moment with one month horizon, see for instance Bollerslev, Tauchen, and Zhou (2009). In the robustness analysis in Section 4.5, we use an alternative approach to compute the risk-neutral variance, following Bakshi and Madan (2000).

## 4.2  Estimating the Physical Second Moment

The literature contains a large number of models for estimating physical variance. In our benchmark analysis, we use a simple and robust implementation of the heterogeneous autoregressive model (HAR) of Corsi (2009), defined as follows

$$V_{t+1,t+K}^m = \phi_0 + \phi_1 V_{t-1,t}^m + \phi_2 V_{t-4,t}^m + \phi_3 V_{t-20,t+K}^m + \varepsilon_{V,t+K}, \tag{22}$$

---

[9]The theoretical results in Section 2 are based on the uncentered moments. Throughout our empirical work we use both centered and uncentered moments. For instance, the VIX is an estimate of the centered second moment. It is well known that centering does not impact estimates of second and third moments much, and we verified that this is indeed the case here.

where $V^m_{t+1,t+K}$ denotes the market index $K$-days ahead integrated variance. In the previous equation, the variance terms satisfy

$$V^m_{s,s+\tau} = V^m_s + V^m_{s+1} + \dots + V^m_{s+\tau}, \tag{23}$$

with the daily variance given as in Rogers and Satchell (1991) by

$$V^m_t = \ln(S^{High}_t/S^{Open}_t)\left[\ln(S^{High}_t/S^{Open}_t) - \ln(S^{Close}_t/S^{Open}_t)\right] \tag{24}$$

$$+ \ln(S^{Low}_t/S^{Open}_t)\left[\ln(S^{Low}_t/S^{Open}_t) - \ln(S^{Close}_t/S^{Open}_t)\right], \tag{25}$$

where $S^{Close}_t$ ($S^{Open}_t$) is the close (open) price of the market index, measured by the S&P 500, and $S^{High}_t$ ($S^{Low}_t$) denotes the market index highest (lowest) price on day $t$. We estimate the HAR model using OLS and a recursive ten-year window. To ensure consistency with our measure of the risk-neutral variance, we generate the one-month forecasts of the physical variance $\hat{V}^m_{t+1,t+30}$ at the end of every month. The HAR model in (22)-(24) parsimoniously allows for a highly persistent dynamic in volatility and employs the intraday information available in our relatively long historical sample. For related applications of high-low information in dynamic volatility models, see Azalideh, Brandt and Diebold (2002), Chou (2005), and Brandt and Jones (2006).[10]

In the robustness analysis in Section 4.5, we use several alternative approaches to estimate the physical variance. We use a simple autoregressive model on realized variances, the NGARCH model of Engle and Ng (1993), and the Heston (1993) stochastic volatility model. For each of these models, we also use a recursive ten-year window.

## 4.3 Option-Based Estimates of the Price of Co-Skewness Risk

Using our benchmark HAR estimate of the physical second moment, and our benchmark VIX risk-neutral second moment, the estimated price of co-skewness risk for month $t$ is now simply

$$\widehat{\lambda}^{COSK}_t = \widehat{E}^P_t(R^2_{m,t+1}) - \widehat{E}^Q_t(R^2_{m,t+1}).$$

Table 3 reports descriptive statistics for the estimates of the moments and the price of risk. Figure 3 depicts the time series of the price of co-skewness with the corresponding estimated

---

[10] Corsi (2009) and subsequent HAR papers typically rely on high-frequency intraday returns to compute daily variance proxies. However, high-frequency returns are not readily available in the beginning of our sample period.

physical and risk-neutral moments required to compute these prices. The plots exhibit spikes surrounding the 1987 stock market crash, the 1998 LTCM collapse, the WorldCom bankruptcy in 2002, and the subprime crisis. These spikes occur for both the risk-neutral and the physical moments, but the spikes in the physical variance are relatively smaller than the risk-neutral spikes except for the case of the subprime crisis. This may be partly due to the model we use for the physical variance. Other approaches for modeling the physical variance in some cases yield larger spikes, but they do not affect our results for cross-sectional pricing. We discuss this further in Section 4.5 below.

Table 3 and Figure 3 indicate that the co-skewness price of risk is negative for almost all months. On average it is equal to −0.271. This negative sign is consistent with theory. The regression-based estimates in Section 3 are also most often negative, and several existing empirical studies document a negative price of co-skewness risk as well, see for instance KL (1976) and HS (2000). However, there are some very important differences between our empirical results and regression-based estimates.

First, our newly proposed estimate of the price of co-skewness risk in Table 3, which is equal to −0.271, is much larger (in absolute value) than any of the estimates obtained using the regression approach in Section 3. This of course does not necessarily mean that our estimate is superior; in order to demonstrate that we have to show that the larger estimate leads to improved fit. We address this in Section 4.4 below.

Second, it is interesting to compare the time-series of conditional estimates in Figure 3 with the time-series for the regression-based estimates in Figure 2. The monthly estimates in Figure 3 are almost all negative, and the difference with Figure 2 is striking. This of course also explains why the negative average estimate of −0.271 for our approach is so much larger (in absolute value), because the negative estimates are not cancelled out by positive estimates in other months.

This comparison between the time series in Figures 3 and 2 must of course be interpreted with some caution. Existing studies report averages of the price of risk over several years. They estimate prices of risk using a two-pass Fama-MacBeth (1973) setup and report the average estimates of the month-by-month cross-sectional regressions, rather than the time series in Figure 2. Indeed, it can be argued that the focus of the Fama-MacBeth approach is to obtain estimates of the price of risk by averaging the time-series in Figure 2, and therefore the time series itself is not meaningful. The month-by-month estimates of the price of risk may not have the theoretically expected negative sign, but this does not invalidate the unconditional estimate.

From this perspective, what is truly remarkable about our new results in Figure 3 is that we have genuinely *conditional* month-by-month estimates of the price of risk that have the theoretically expected sign in almost every month. The regression-based approach obviously does not provide us with such results. Moreover, while there is no guarantee that these negative estimates for the price of co-skewness risk will continue to obtain in the future, we know that implied variances usually exceed historical variances. Because of this stylized fact, our approach is more likely to yield plausible estimates of co-skewness risk.

In summary, a comparison of our newly proposed estimates of the price of co-skewness risk with regression-based estimates yields three important conclusions. First, regression-based estimates critically depend on the test assets used in estimation, whereas our approach is by design independent of the test assets. Second, our unconditional estimate of the price of co-skewness risk is $-0.271$ and indicates a role for co-skewness that is much larger in magnitude compared to regression-based approaches. Third, we consistently obtain negative estimates of the price of conditional co-skewness risk in our approach, which does not obtain with regression-based methods.

We therefore conclude that our approach is economically appealing. To show that it improves on regression-based estimates, we have to demonstrate that it leads to a better fit. This is the subject to which we now turn.

## 4.4 Comparing Out-of-Sample Model Fit

When using regression-based methods, the cross-sectional or Fama-MacBeth 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, Table 2 reports on model performance using the $R$-square. Even though there are many other related evaluation criteria, in the overwhelming majority of cases these criteria are in-sample as in Table 2. Table 2 highlights a common drawback of such in-sample comparisons in which models with more factors often lead to a better fit.

It is important to note that in our proposed approach betas and loadings are constructed 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 is 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 can be estimated using the historical market risk premium, and the cross-sectional fit of the CAPM can 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 compare our models using a genuine out-of-sample approach for all models. Out-of-sample comparisons of cross-sectional models is becoming increasingly popular, see for instance Simin (2008) and Ferson, Nallareddy, and Xie (2012).

Simin (2008) and Ferson, Nallareddy, and Xie (2012) use root mean squared errors to compare cross-sectional models out of sample. We use the out-of-sample $R$-square forecast evaluation criterion 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 $R$-squares lead to the same model ranking as root mean squared errors but their magnitudes are easier to interpret. The out-of-sample $R^2_{j,OS}$ for a security $j$ is defined by

$$
R^2_{j,OS} = 1 - \frac{\sum_t \left( R_{j,t+1} - \widehat{R}^{Model}_{j,t+1} \right)^2}{\sum_t \left( R_{j,t+1} - \overline{R}_{j,t-59:t} \right)^2}
\tag{26}
$$

where $\overline{R}_{j,t-59:t} = \frac{1}{60} \sum_{s=t-59}^{t} R_{j,s}$ is the benchmark forecast constructed as the average of the past 60 monthly returns. We report the average $R^2_{OS}$ across portfolios for each model.

Note that this out-of-sample $R$-square uses the historical average return on the test portfolio as a benchmark. If a candidate model performs as well as the historical average return on the test portfolio, the resulting $R$-square will be zero. $R$-squares can be negative for models that do not perform well in out-of-sample 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 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 risk premia are estimated using cross-sectional regressions. The models that use cross-sectional regressions to

21

estimate the risk premia are the model with market covariance risk (the $CAPM$), the model with market covariance and co-skewness risk ($CAPM + COSK$), and the Fama-French three-factor model ($FF$). The specifications based on historical risk premia are: $CAPM$, $COSK$, and $CAPM + COSK$. We also include a hybrid approach, labeled $CSCAPM + COSK$, where the market risk premium is estimated using a cross-sectional regression.

To provide more intuition, consider 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

$$\widehat{R}_{j,t+1}^{CAPM} = \widehat{\lambda}_t^{MKT} \widehat{\beta}_{j,t}^{MKT} \tag{27}$$

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 estimates the price of risk using a cross-sectional regression:

$$\mathbf{R}_t = \boldsymbol{\lambda}_t^{MKT} \; \widehat{\boldsymbol{\beta}}_{t-1}^{MKT} + \mathbf{u}_t, \tag{28}$$

where the vectors $\mathbf{R}$ and $\widehat{\boldsymbol{\beta}}$ contain the cross-section of portfolio returns and $CAPM$ betas respectively. 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 in Figure 2. To obtain better out-of-sample competitors for our estimators of co-skewness risk that are based on historical risk premia, we therefore use 60-month averages of the cross-sectional $\boldsymbol{\lambda}_t$. Arguably, this approach is also more in line with the conventional (in-sample) implementation of Fama-MacBeth regressions.

Table 4 presents the results. Recall that a positive out-of-sample $R$-square means that the model forecasts better than the historical average return on the asset. The performance of our newly proposed co-skewness measure $COSK$ in the second column of the top four rows is promising. It yields a positive $R$-square for all four sets of test portfolios. The out-of-sample performance of the other models is mixed. Arguably the best competitor is the regression-based implementation of the $CAPM$, but this model does poorly for the twenty-five size and book-to-market portfolios. The out-of-sample performance of the Fama-French model is disappointing. It may seem surprising that the $FF$ model performs so poorly for

the case of the 25 size and book-to-market portfolios, but note that the $FF$ model is not typically evaluated in a out-of-sample setting.

It is also interesting to investigate the forecasting performance of the newly proposed co-skewness measure during different market conditions. Table 5 presents the out-of-sample $R$-squares of the models studied in Table 4 for good and bad times separately. Bad (Good) times correspond to years with negative (positive) market returns. A comparison of Panel A and Panel B indicates that the forecasting performance is better during bad economic times for all models. The $COSK$ model performs exceptionally well in bad times. The model's average out-of-sample $R$-square ranges from 3.38% to 4.83% across test assets, which is impressive. While the $COSK$ model has a negative out-of-sample $R$-square in good times, its performance is roughly similar that of other models such as the $CAPM$ implemented using historical risk premia.

It is important to keep in mind 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 then $R$-squares of $1 - 2\%$ are the exception rather than the rule, with many candidate forecasts yielding negative $R$-squares, see Campbell and Thompson (2008), Welch and Goyal (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 noteworthy, especially because forecasting the cross-section of returns is arguably even harder than time series forecasting.

## 4.5   Robustness

We now report on several robustness exercises, using alternative measures of conditional physical and risk-neutral second moments.

We used the VIX as our measure of the risk-neutral second moment in our benchmark results. In the robustness analysis we use an alternative following Bakshi and Madan (2000) and Bakshi, Kapadia, and Madan (2003). This approach is based on a continuum of out-of-the money call and put options which is approximated using cubic spline interpolation techniques.

Let $S_t$ denote the value of the market index and $R_{m,t+\tau} = \ln S_{t+\tau} - \ln S_t$ its return over

the horizon $\tau$. We can get the risk-neutral second moment via

$$E_t^Q \left[ R_{m,t+\tau}^2 \right] = e^{r\tau} \int_{S_t}^{\infty} \frac{2 \left( 1 - \ln \left[ K/S_t \right] \right)}{K^2} C_t \left( \tau, K \right) dK \tag{29}$$
$$+ e^{r\tau} \int_0^{S_t} \frac{2 \left( 1 + \ln \left[ S_t/K \right] \right)}{K^2} P_t \left( \tau, K \right) dK.$$

where $C_t \left( \tau, K \right)$ and $P_t \left( \tau, K \right)$ are call and put options quoted at time $t$ with maturity $\tau$ and strike price $K$. See Appendix A for more details on the implementation and the option data used.

We investigate three alternative approaches for modeling the conditional physical variance. We first consider a simple autoregressive model on realized variances. The one-step ahead forecast of the physical second moment is estimated from the following monthly regression

$$V_t^m = a_0 + a_1 V_{t-1}^m + u_{V,t}, \tag{30}$$

where $V_t^m = \sum_{d \in t} R_{m,d,t}^2$, $R_{m,d,t}$ denotes the daily market index return in day $d$ of month $t$, and $u_{V,t}$ is the variance innovation.

In addition to the autoregressive model we also use an NGARCH model (Engle and Ng, 1993) to estimate the physical variance

$$R_{m,t} = \sqrt{V_t^m} z_t \qquad z_t \sim N(0,1), \tag{31}$$
$$V_t^m = a_0 + b_0 V_{t-1}^m \left( z_{t-1} - d_0 \right)^2 + c_0 V_{t-1}^m. \tag{32}$$

The $T$-days ahead aggregate volatility forecast can be computed as follows

$$E_t[R_{m,t+1:t+T}^2] = T V_0^m + (V_{t+1}^m - h_0^2) \frac{1 - (b_0 + c_0 + b_0 d_0^2)^T}{1 - b_0 - c_0 - b_0 d_0^2}, \tag{33}$$

where $V_0^m = \frac{a_0}{1 - b_0 - c_0 - b_0 d_0^2}$. Finally we also use the Heston (1993) stochastic volatility model in which the market index return follows

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{V_t^m} dW_{S,t}, \tag{34}$$

and the instantaneous variance dynamic is

$$dV_t^m = \kappa(\theta - V_t^m)dt + \eta \sqrt{V_t^m} dW_{V,t}, \tag{35}$$

24

where $W_{S,t}$ and $W_{V,t}$ are two correlated Brownian motion processes with $dW_{V,t}dW_{S,t} = \rho dt$. We estimate this model using maximum likelihood and the particle filter.

Table 6 presents the results. Panel A contains the estimates of the price of risk obtained using the different approaches. Panel B contains the out-of-sample $R$-squares $R^2_{OS}$ for the $COSK$ model.

The estimates of the price of risk in Panel A vary between $-0.123$ and $-0.316$. Recall that our benchmark estimate in Table 3 was $-0.271$. These estimates are quite similar and they are all larger (in absolute value) than the cross-sectional estimates in Table 2. The out-of-sample $R$-squares in Panel B are positive in twenty-six out of twenty-eight cases, which is quite impressive when compared with the models in Table 4. We conclude that our newly proposed estimates of the price of risk are rather robust across different empirical implementations, and that the resulting out-of-sample performance is much better than that of regression-based implementations of models with co-skewness risk, as well as competing factor model specifications.

# 5    The Price of Co-Kurtosis Risk

We now provide estimates of the price of co-kurtosis risk using the estimator presented in Proposition 2. To estimate the price of co-kurtosis risk, we need to estimate the third conditional moment under the risk-neutral measure $E_t^Q(R^3_{m,t+1})$ and under the physical measure $E_t^P(R^3_{m,t+1})$.

Estimating the third moment is much harder than estimating the second moment, especially in the case of the physical third moment. We first discuss our benchmark implementation, which is as simple as possible to minimize the impact of modeling choices. Subsequently we discuss more sophisticated approaches.

## 5.1    Estimating Risk-Neutral and Physical Third Moments

Evaluating the pricing of co-kurtosis risk is arguably most meaningful if lower moment risk is also considered. The empirical results in this section thus not only require estimates of risk-neutral and physical third moments, but also of risk-neutral and physical second moments. For the physical second moment we again use the benchmark HAR model in equation (22)-(24), and for the risk-neutral second moment we use the VIX benchmark.

With respect to the modeling of the third moment, it is well known that capturing the

time-variation in the physical third moment is extremely difficult, see for instance Jondeau and Rockinger (2003). It is also well known that for index returns, the risk-neutral third moment is on average much larger (in absolute value) than the physical third moment, see for instance Bakshi and Kapadia (2003).

Given these stylized facts, we proceed as follows. We estimate the risk-neutral third moment using the method of Bakshi and Madan (2000). We implement this approach using data on S&P500 index options from OptionMetrics for the period January 1996 to December 2012. Note that this sample period is different from the one used in Section 4. As in equation (29) above let $S_t$ denote the value of the market index and $R_{m,t+\tau} = \ln S_{t+\tau} - \ln S_t$ its return over the horizon $\tau$. We can get the option-implied third moment via

$$
\begin{aligned}
E_t^Q\left[R_{m,t+\tau}^3\right] \quad = \quad & e^{r\tau}\int_{S_t}^{\infty}\frac{6\ln\left[K/S_t\right] - 3\left(\ln\left[K/S_t\right]\right)^2}{K^2}C_t\left(\tau, K\right)dK \\
& -e^{r\tau}\int_0^{S_t}\frac{6\ln\left[S_t/K\right] + 3\left(\ln\left[S_t/K\right]\right)^2}{K^2}P_t\left(\tau, K\right)dK.
\end{aligned}
\tag{36}
$$

where $C_t\left(\tau, K\right)$ and $P_t\left(\tau, K\right)$ are call and put options quoted at time $t$ with maturity $\tau$ and strike price $K$. See Appendix A for more details on the implementation and the data.

With respect to the physical third moment, we simply set it to zero throughout the sample in our benchmark implementation. Confirming existing results, the unconditional third moment estimate for monthly S&P500 returns during 1996-2012 is not statistically different from zero at conventional confidence levels, and moreover it is very small compared to the estimates of risk-neutral moments in our sample. Setting the physical third moment to zero may therefore be preferable to using noisy estimates. In Section 5.3, we explore the results using several alternative estimates for the physical third moments: unconditional sample moments and estimates provided by a dynamic model of the physical third moment.

## 5.2   Estimates of The Price of Co-Kurtosis Risk

The price of co-kurtosis risk for month $t$ can now simply be computed as

$$
\widehat{\lambda}_t^{COKU} = \widehat{E}_t^P(R_{m,t+1}^3) - \widehat{E}_t^Q(R_{m,t+1}^3).
$$

For our benchmark implementation this gives

$$
\widehat{\lambda}_t^{COKU} = -\widehat{E}_t^Q(R_{m,t+1}^3).
$$

26

Figure 4 depicts the time series of the price of co-kurtosis risk for our benchmark implementation, which is simply the negative of the risk-neutral third moment. Consistent with theory, the price of co-kurtosis risk in Figure 4 is positive throughout the period. It is equal to 0.022 on average.[11] Existing empirical studies have also documented positive prices of co-kurtosis risk. See for instance Ang, Chen, and Xing (2006), who find that stocks with higher co-kurtosis earn higher returns.

Figures 5 and 6 report on estimates of co-kurtosis risk obtained using Fama-MacBeth regressions. Figure 5 indicates that the month-by-month estimates of the price of co-kurtosis risk vary considerably over time, and that they are often negative. Compared to the regression-based time series of the prices of co-skewness risk in Figure 2, the time series of the prices of co-kurtosis risk are less correlated across test assets. When averaging over time, the estimate is significantly negative for the twenty-five size and momentum portfolios. This is also the case for the univariate regressions in Panel A of Figure 6. However, Panel B of Figure 6 indicates that this may be due to the relatively short sample period. When using the longer 1966-2012 time period, all four estimates of co-skewness risk are positive, although not always statistically significant.

For the 1996-2012 sample period in Panel A of Figure 6, only one set of test portfolios yields a statistically significant positive result, the twenty-five size and short term reversal portfolios. The resulting estimate of the price of co-kurtosis risk is 0.020. The estimates obtained for the 1966-2012 sample period in Panel B of Figure 6 are of the same order of magnitude, but somewhat larger. We conclude that our new estimates of the price of co-kurtosis risk are rather similar to regression-based estimates, which contrasts with our findings on the price of co-skewness risk.

Table 7 presents out-of-sample $R$-squares using these estimates of the prices of co-skewness and co-kurtosis risk, and compares the resulting fit with the fit of regression-based approaches. As in Table 4, the price of risk is estimated using cross-sectional regressions or historical risk premia. When using historical risk premia, we provide out-of-sample predictions for the $CAPM$, the model with a co-skewness premium $COSK$, the model with a co-kurtosis premium $COKU$, and the model with market and co-skewness factors and market and co-kurtosis factors, $CAPM + COSK$ and $CAPM + COKU$ respectively. When using cross-sectional regressions, we provide predictions for the $CAPM$, $CAPM + COSK$, $CAPM + COKU$, and the Fama-French three-factor model $FF$. We also use hybrid models

---

[11]Recall that for expositional convenience these estimates of the moments and the prices of co-skewness and co-kurtosis risk are reported in percentage terms, i.e. multiplied by 100.

with regression-based co-variance premium and historical co-skewness and co-kurtosis premiums, $CSCAPM + COSK$ and $CSCAPM + COKU$. Furthermore, we use the HAR+VIX benchmark model in order to obtain the time-series of the price of co-skewness risk used in the $COSK$, $CAPM + COSK$, and $CSCAPM + COSK$ specifications.

Note that the sample period is different from the one used in Section 4.3. However, the resulting estimates of the price of co-skewness risk are similar to the ones obtained in Section 4.3, and it is therefore not surprising that the resulting $R$-squares are similar to the ones in Table 4.

The model with co-kurtosis risk performs well when using price of risk estimated from historical risk premia. It has large, positive out-of-sample $R$-squares for the 25 size and book-to-market portfolios and for the 25 size and short-term reversal portfolios. For the 25 size and momentum portfolios, the $R$-square is positive but not large and for the 25 size and long-term reversal it is slightly negative. These option-based results for the co-kurtosis model compare very favorably with those for the other models in Table 7.

## 5.3   The Dynamics of the Physical Third Moment

It is well known that modeling the conditional third moment is challenging, partly because it is much less persistent than the second moment–particularly so at the monthly frequency. Our own empirical implementation confirmed these challenges. Together with the knowledge that the risk-neutral third moment is much larger than the physical third moment, this motivated us to set the physical third moment equal to zero in our empirical implementation. We now investigate if improvements can be made through alternative modeling assumptions.

Table 8 reports on the $COKU$ model using three alternative estimates for the physical third moment: first, a constant third moment computed using daily data; second, a constant third moment computed using monthly data; and finally a fully dynamic physical third moment. To implement the dynamic physical third moment, we use a version of the dynamic moment model in Jondeau and Rockinger (2003) described in Appendix B. Note that this model includes the specification of a dynamic second moment. Our implementation is close to the model Jondeau and Rockinger (2003) refer to as Model 2, which is among the more parsimonious models they consider and which is sufficiently richly parameterized for our purposes.

The results in Panel A of Table 8 indicate that the resulting estimates of the price of co-kurtosis risk are on average similar to the estimate of 0.022 obtained using a zero physical

third moment. Note from Panel B that the out-of-sample performance of the co-kurtosis model when allowing for a dynamic physical third moment is substantially worse than when using zero or a constant for the physical estimate. This may suggest that our estimates of the conditional physical third moment are too noisy, which may affect the cross-sectional out-of-sample fit. We therefore suggest simply setting the physical third moment equal to zero as we do in Table 7 until methods become available that make it possible to estimate the dynamic third moment more accurately and precisely.

# 6   Conclusion

We propose an alternative strategy for estimating the price of possibly nonlinear exposures to market risk that avoids the errors inherent in the cross-sectional regression approach. The key difference between our approach and existing methods is that we explicitly impose the consistent pricing restrictions on both stocks and contingent claims. We study two important applications of our general approach: The price of co-skewness risk in our framework corresponds to the spread between the physical and the risk-neutral second moment. 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 almost every month, and the price of co-kurtosis risk has the theoretically expected positive sign in most months. In contrast, the prices of risk obtained using regression-based approaches do not always have the theoretically anticipated signs even on average. Our approach also provides genuinely conditional estimates of the price of risk at monthly or even higher frequencies. When using a regression-based approach, monthly estimates are available, but they are very imprecise, and they are therefore usually averaged over a large number of months. 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 competing factor models such as the CAPM and the Fama-French model.

Several 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, alternative computations of the physical and risk-neutral moments are needed, especially for the case of

co-kurtosis. It may prove interesting to use the CBOE SKEW index, which, like the VIX, is readily available. 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 estimated betas we use in the analysis may be noisy. 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. Finally, it would be useful to reliably assess the statistical significance of the price of co-skewness and co-kurtosis risk that takes into account the uncertainty in the various steps involved in the computation.

# Appendix A: Extracting Option-Implied Moments

We estimate the risk-neutral second and third moments using the method of Bakshi and Madan (2000). We implement this approach using data on S&P500 index options from OptionMetrics for the period January 1996 to December 2012. 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 available range, we simply use the implied volatility of the lowest or highest available strike price.

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.

Let $S_t$ denote the value of the market index and $R_{m,t+\tau} = \ln S_{t+\tau} - \ln S_t$ its return over the horizon $\tau$. We can get the risk-neutral second moment via

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

where $C_t\left(\tau, K\right)$ and $P_t\left(\tau, K\right)$ are call and put options quoted at time $t$ with maturity $\tau$ and strike price $K$. We can get the option-implied third moment via

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

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 continuum 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.

# Appendix B: Modeling Dynamic Physical Third Moments

We implement the Jondeau and Rockinger (2003) model using monthly data. The model is given by

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

where $R_{m,t}$ is the return on the market in month $m$, $GT$ denotes the generalized student-t distribution, and where the higher-moment dynamics are modeled via

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

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. We do so for two reasons. First, the mean is difficult to estimate reliably. Second, and maybe most importantly, the first moment of the market returns has little impact on 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

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

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

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

and

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

where

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

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

# References

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

[2] Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006), The Cross-Section of Volatility and Expected Returns, *Journal of Finance*, 61, 259–299.

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

[4] Azalideh, S., M. Brandt, and F. Diebold (2002), Range-Based Estimation of Stochastic Volatility Models, *Journal of Finance*, 57, 1047–1091.

[5] Bakshi, G., and N. Kapadia (2003), Delta-Hedged Gains and the Negative Market Volatility Risk Premium, *Review of Financial Studies*, 16, 527–566.

[6] 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.

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

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

[9] 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.

[10] 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.

[11] 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.

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

[13] Bansal, R., R. Dittmar, and C. Lundblad (2005), Consumption, Dividends, and the Cross-Section of Equity Returns, *Journal of Finance*, 60, 1639–1672.

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

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

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

[17] Bollerslev, T., J. Marrone, L. Xu, and H. Zhou (2014), Predicting Stock Returns with Variance Risk Premia: Statistical Inference and International Evidence, *Journal of Financial and Quantitative Analysis*, 49, 633–661.

[18] Brandt, M., and C. Jones (2006), Volatility Forecasting with Range-Based EGARCH Models, *Journal of Business and Economic Statistics*, 24, 470–486.

[19] Breeden, D., M. Gibbons, and R. Litzenberger (1989), Empirical Tests of the Consumption-Oriented CAPM, *Journal of Finance*, 44, 231–262.

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

[21] 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.

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

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

[24] 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.

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

[26] Chou, R. (2005), Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range, *Journal of Money, Credit and Banking*, 37, 561–582.

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

[28] Corsi, F. (2009), A Simple Approximate Long Memory Model of Realized Volatility, *Journal of Financial Econometrics*, 7, 174–196.

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

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

[31] Dittmar, R.F. and C. Lundblad (2015), Firm Characteristics, Consumption Risk, and Firm-Level Risk Exposures, Working Paper, University of Michigan.

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

[33] Engle, R., and V. Ng (1993), Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, 48, 1749–1778.

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

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

[36] 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.

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

[38] Gospodinov, N., R. Kan, and C. Robotti (2014), Misspecication-Robust Inference in Linear Asset Pricing Models with Irrelevant Risk Factors, *Review of Financial Studies*, 27, 2139–2170.

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

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

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

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

[43] Jagannathan, R., and Z. Wang (1998), An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression, *Journal of Finance*, 53, 1285–1309.

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

[45] Kan, R., and C. Zhang (1999), Two-Pass Tests of Asset Pricing Models with Useless Factors, *Journal of Finance*, 54, 204–235.

[46] Kleibergen, F. (2009), Tests of Risk Premia in Linear Factor Models, *Journal of Econometrics*, 149, 149–173.

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

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

[49] Lettau, M., and S. Ludvigson (2001), Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, *Journal of Political Economy*, 109, 1238–1287.

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

[51] 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.

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

[53] Mankiw, G., and M. Shapiro (1986), Risk and Return: Consumption Beta versus Market Beta, *Review of Economics and Statistics*, 68, 452–459.

[54] 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.

[55] 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.

[56] Rogers, L. and S. Satchell (1991), Estimating Variance from High, Low and Closing Prices, *Annals of Applied Probability*, 1, 504–512.

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

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

[59] Schneider, P., C. Wagner, and J. Zechner (2015), Low-Risk Anomalies? Working Paper, Copenhagen Business School.

[60] 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.

[61] Shanken, J. (1992), On the Estimation of Beta Pricing Models, *Review of Financial Studies*, 5, 1–33.

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

[63] Welch, I., and A. Goyal (2008), A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies*, 21, 1455–1508.

# Figure 1: The Cross-Section of Returns and Co-Skewness Betas

We plot average excess returns (monthly, in percentages), $E[R_j] - R_f$, against co-skewness betas, $\beta_j^{COKU}$, for four sets of portfolios. The co-skewness beta, $\beta_j^{COSK}$, is computed from the regression of monthly excess returns on market returns and squared market returns. We consider two periods, 1986-2012 (Panel A), and 1966-2012 (Panel B), and four sets of test portfolios. The $\lambda$s are multiplied by 100 for expositional convenience.

## Panel A. 1986 - 2012

**Panel B. 1966 - 2012**

### 25 Size/Book-to-Market

$$E[R_j] - R_f = \lambda^0 + \lambda^{CSK} \beta_j^{CSK}$$

0.588    -0.168
(10.29)    (-1.92)

### 25 Size/Momentum

$$E[R_j] - R_f = \lambda^0 + \lambda^{CSK} \beta_j^{CSK}$$

0.523    -0.347
(12.40)    (-7.34)

### 25 Size/Short-Term Reversal

$$E[R_j] - R_f = \lambda^0 + \lambda^{CSK} \beta_j^{CSK}$$

0.632    0.074
(8.61)    (0.66)

### 25 Size/Long-Term Reversal

$$E[R_j] - R_f = \lambda^0 + \lambda^{CSK} \beta_j^{CSK}$$

0.677    -0.147
(17.28)    (-1.81)
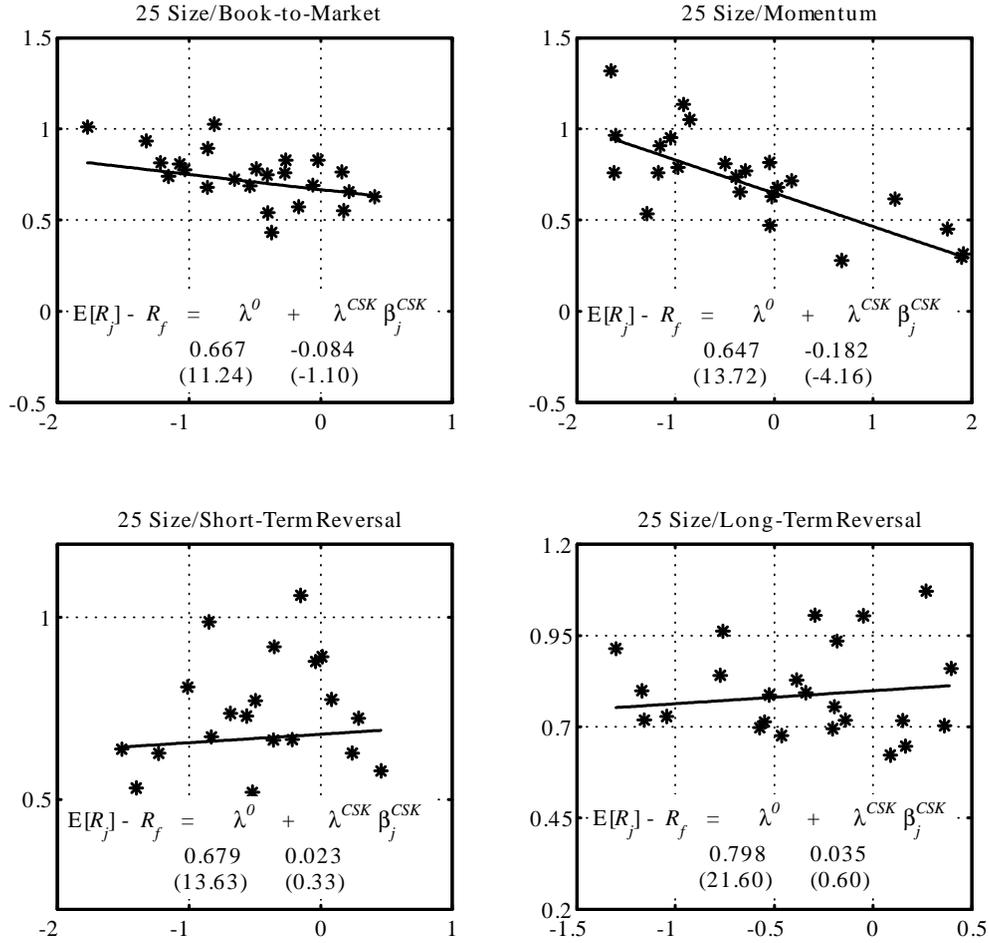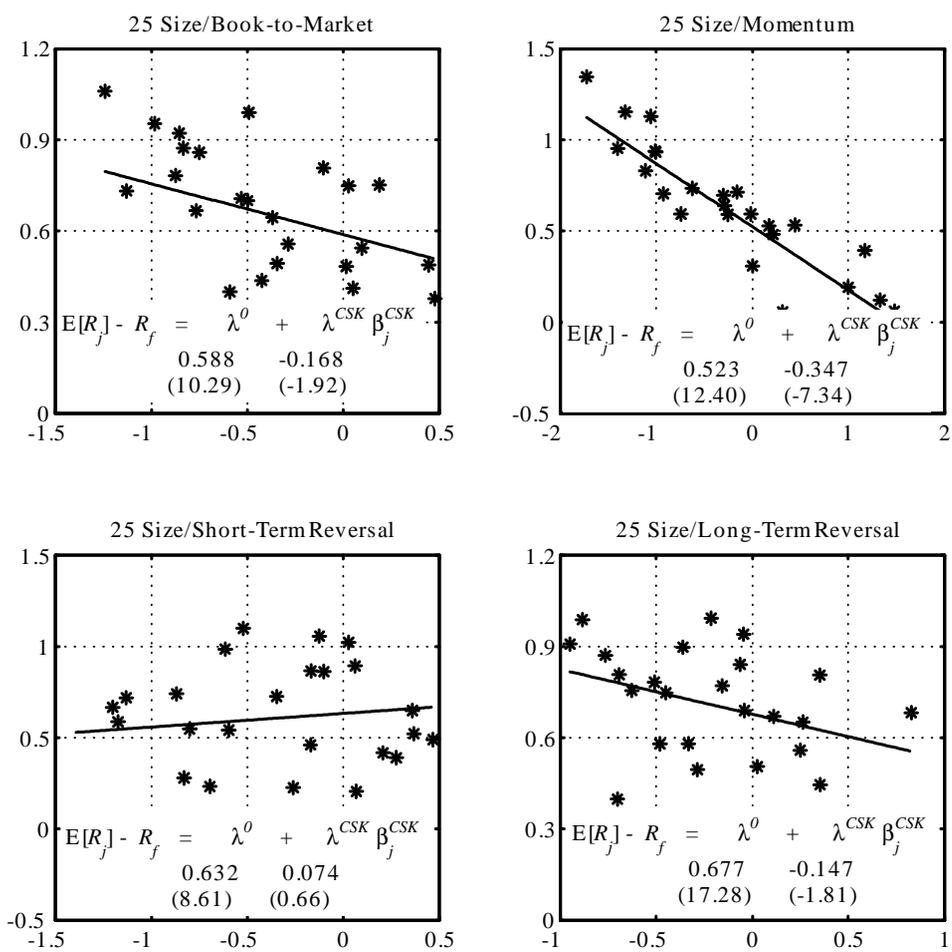
## Figure 2: Regression-Based Estimates of the Price of Co-Skewness Risk

We plot time series of the cross-sectional prices of co-skewness risk, multiplied by 100 for expositional convenience. Each month, we estimate the co-skewness beta using a 60-month rolling window of monthly returns from the following time series regression

$$R_{j,t} - R_{f,t} = \alpha_{j,t} + \beta_{j,t}^{MKT} R_{m,t} + \beta_{j,t}^{COSK} R_{m,t}^2 + \varepsilon_{j,t}.$$

We then run the following cross-sectional regression using the estimated betas and returns for the next month

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

We consider four sets of test portfolios: 25 Size/Book-to-Market, 25 Size/Momentum, 25 Size/Short-Term Reversal and 25 Size/Long-Term Reversal. The sample period is from January 1986 through December 2012.
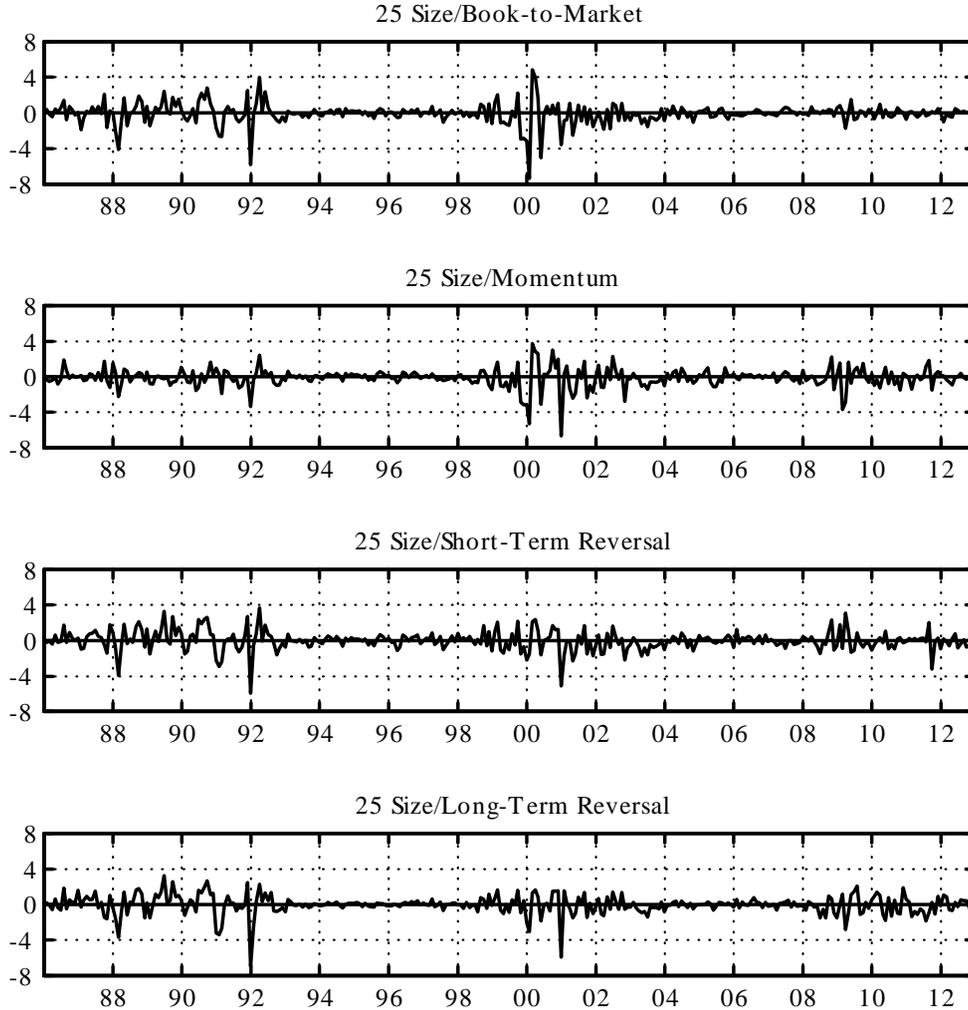
## Figure 3: The Option-Based Price of Co-Skewness Risk

We plot the time series for the conditional physical and risk-neutral second moments (monthly in percentage) and the price of co-skewness risk, multiplied by 100 for expositional convenience. The physical second moment is estimated using an HAR model and the risk-neutral second moment is proxied by the VIX squared. 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 1986 to December 2012.
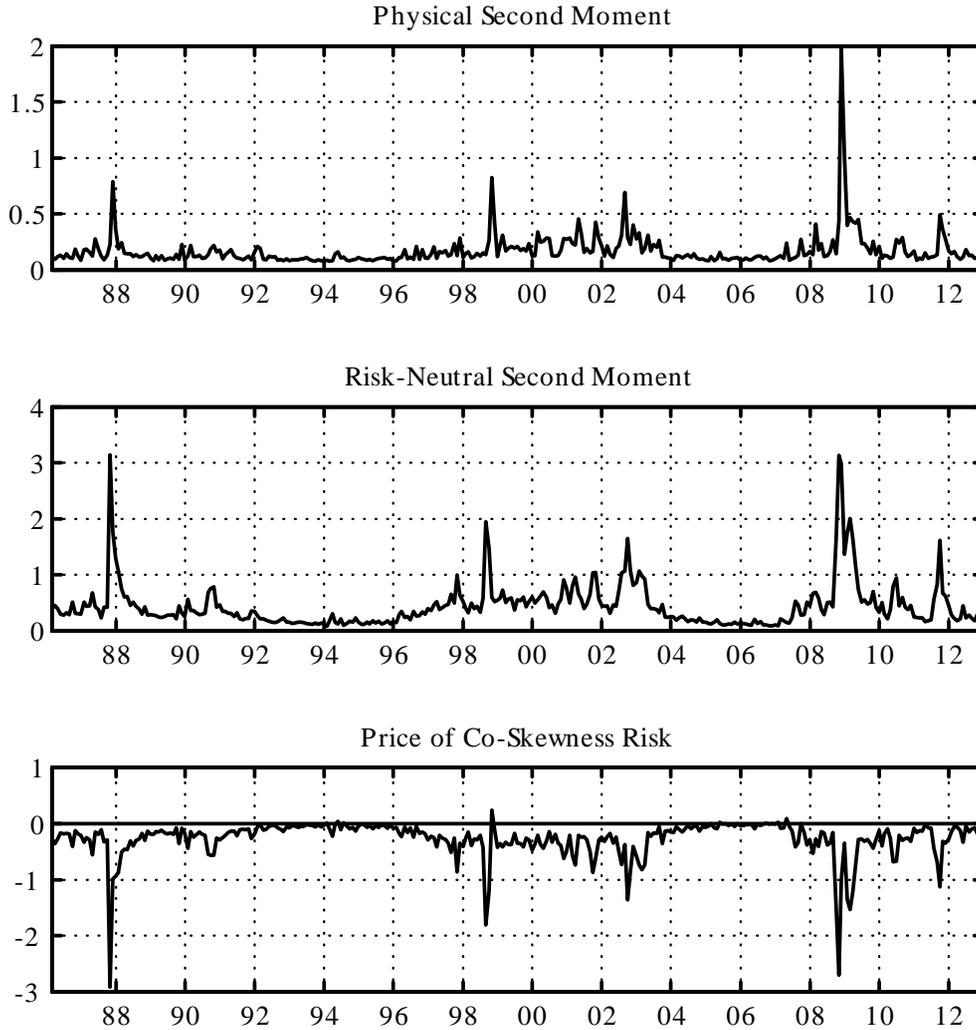
# Figure 4: The Option-Based Price of Co-Kurtosis Risk

We plot the price of co-kurtosis risk, multiplied by 100 for expositional convenience. We report on the benchmark case where the physical third moment is set equal to zero. The risk-neutral moment is estimated using the model-free approach in Bakshi and Madan (2000) and Bakshi, Kapadia, and Madan (2003). The sample period is from January 1996 to December 2012.



Price of Co-Kurtosis Risk
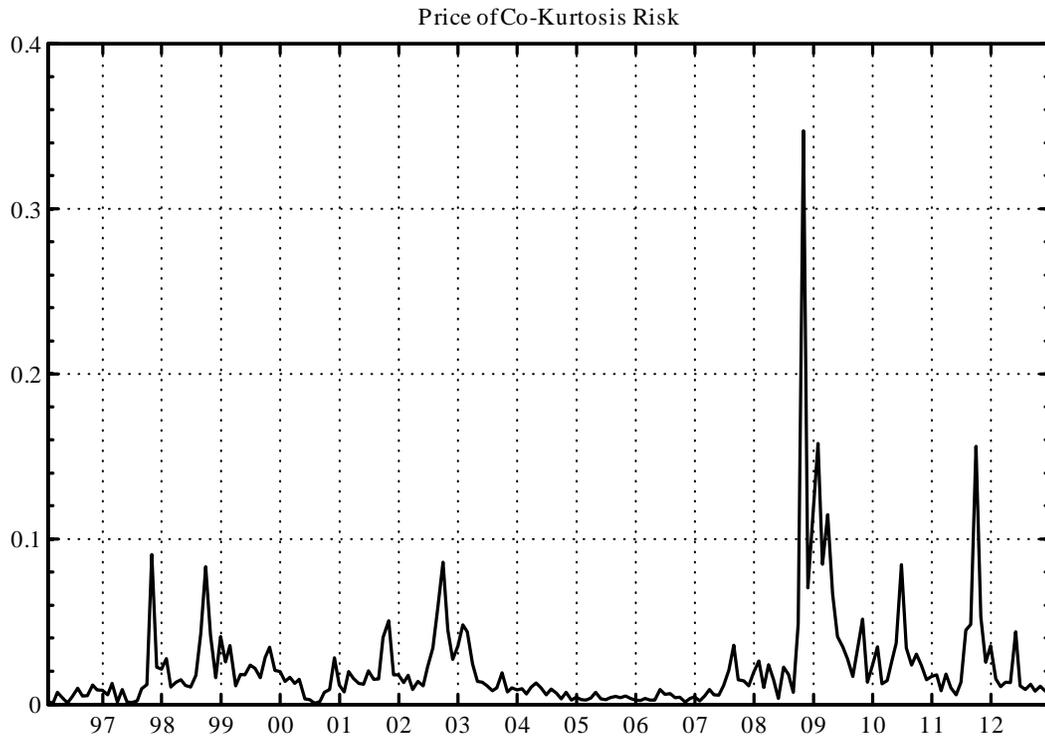
# Figure 5: Regression-Based Estimates of the Price of Co-Kurtosis Risk

We plot the time series for the cross-sectional price of co-kurtosis risk, multiplied by 100 for expositional convenience. Each month, we estimate the co-kurtosis beta using a 60-month rolling window of monthly returns from the following time series regression

$$R_{j,t} - R_{f,t} = \alpha_{j,t} + \beta_{j,t}^{MKT} R_{m,t} + \beta_{j,t}^{COSK} R_{m,t}^2 + \beta_{j,t}^{COKU} R_{m,t}^3 + \varepsilon_{j,t}.$$

We then run the following cross-sectional regression using the estimated betas and returns for the next month

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

We consider four sets of portfolios: 25 Size/Book-to-Market, 25 Size/Momentum, 25 Size/Short-Term Reversal and 25 Size/Long-Term Reversal. The sample period is from January 1996 to December 2012.



44
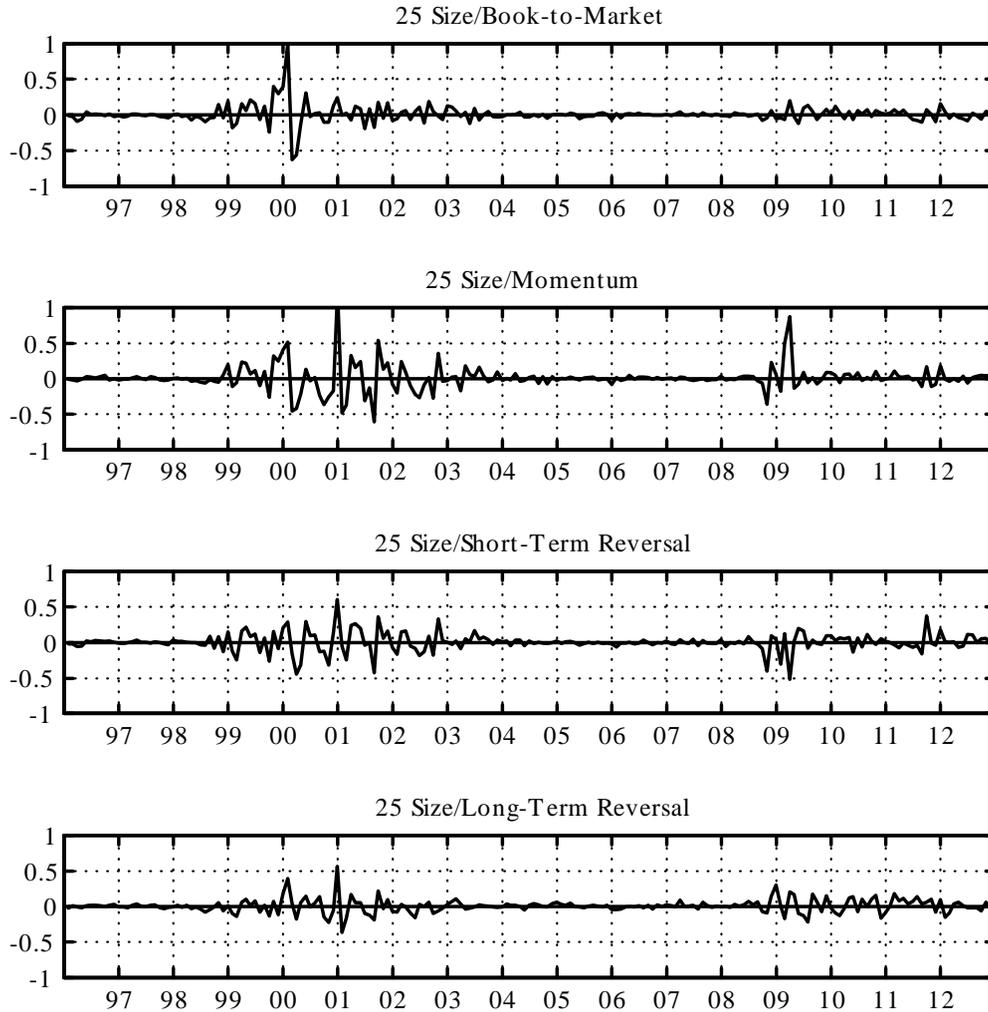
# Figure 6: The Cross-Section of Returns and Co-Kurtosis Betas

We plot average excess returns (in percentage per month), $E[R_j] - R_f$, against co-skewness betas, $\beta_j^{COKU}$, for four sets of portfolios. The co-skewness beta, $\beta_j^{COKU}$, is computed from the regression of monthly excess returns on market returns, squared market returns and cubic market returns. We consider two periods, 1996-2012 (Panel A) and 1966-2012 (Panel B), and four sets of test portfolios. The $\lambda$s are multiplied by 100 for expositional convenience.

### Panel A. 1996 - 2012

### 25 Size/Book-to-Market



$$E[R_j] - R_f = \lambda^0 + \lambda^{COKU} \beta_j^{COKU}$$

0.669    0.030
(15.75)   (1.94)

### 25 Size/Momentum



$$E[R_j] - R_f = \lambda^0 + \lambda^{COKU} \beta_j^{COKU}$$

0.608    0.016
(8.32)   (0.87)

### 25 Size/Short-Term Reversal



$$E[R_j] - R_f = \lambda^0 + \lambda^{COKU} \beta_j^{COKU}$$

0.641    0.057
(11.13)   (2.25)

### 25 Size/Long-Term Reversal



$$E[R_j] - R_f = \lambda^0 + \lambda^{COKU} \beta_j^{COKU}$$

0.661    0.031
(17.46)   (2.57)
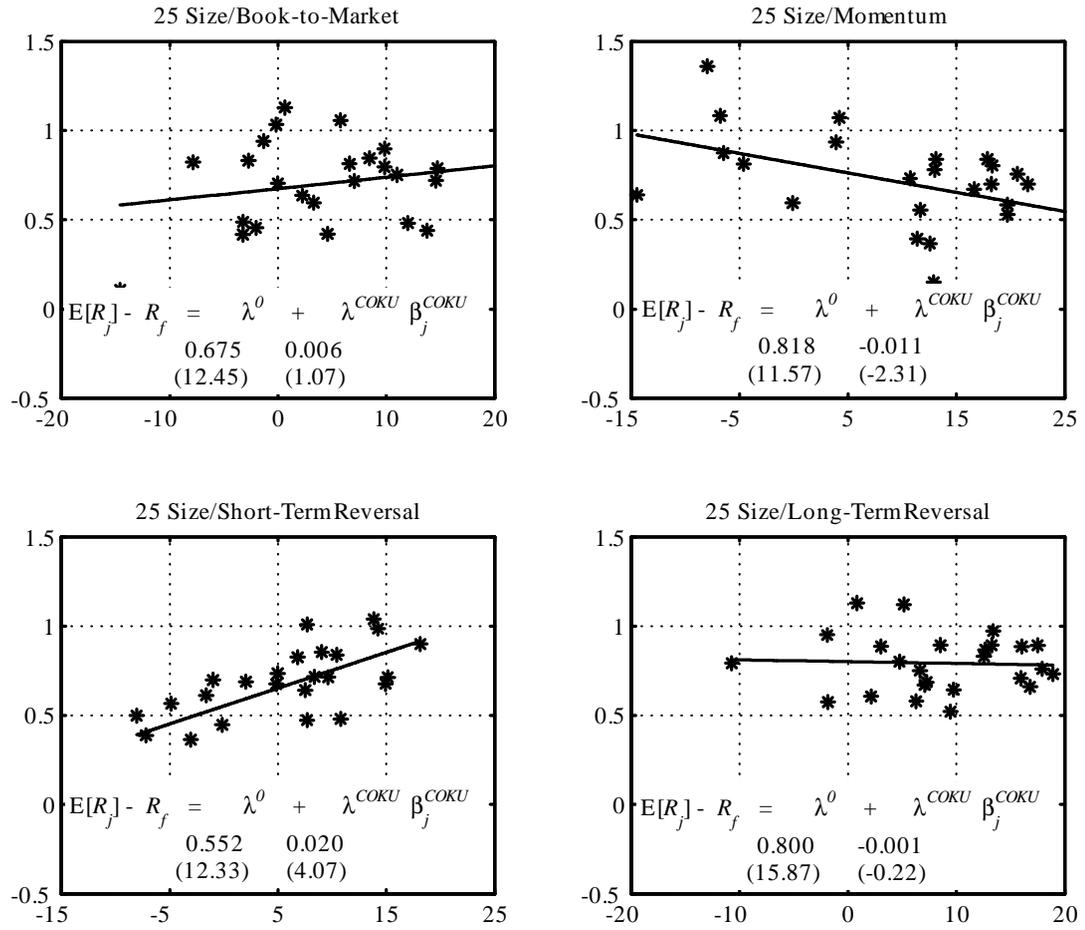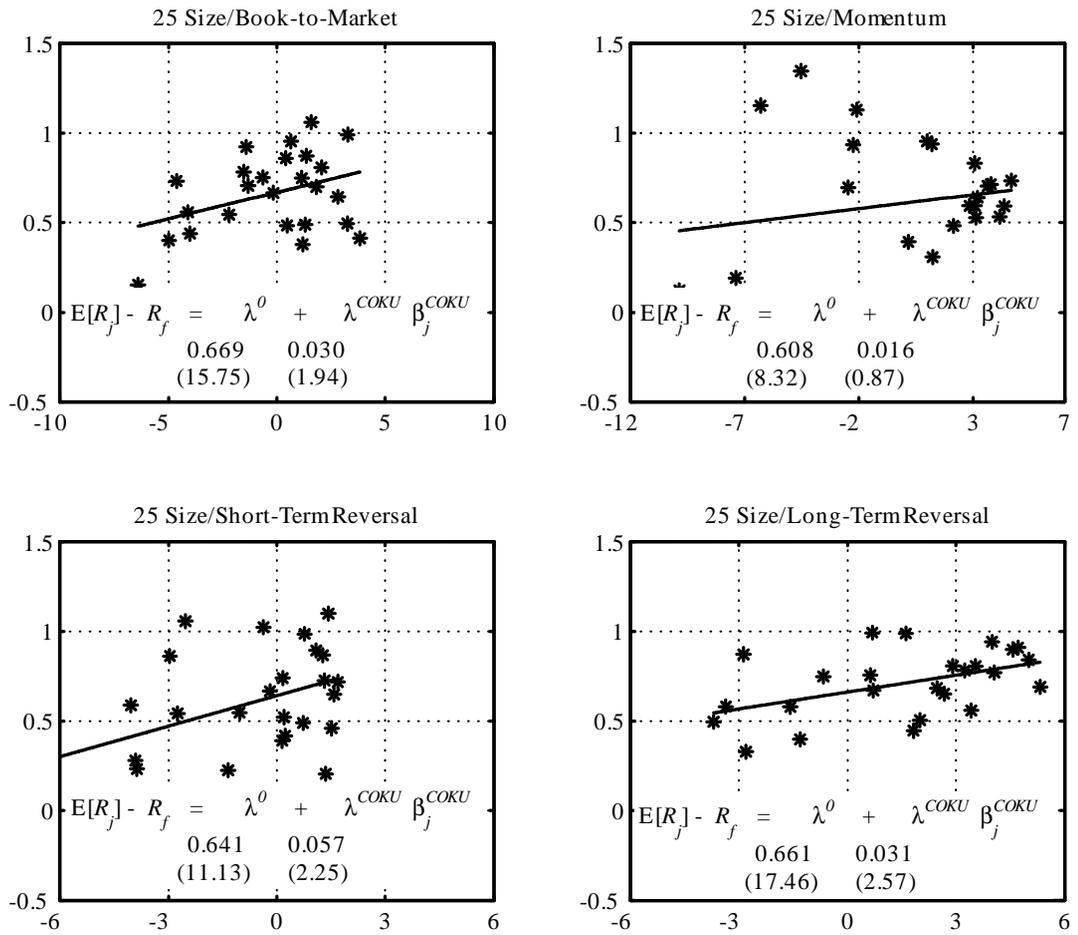
## Table 1: The Price of Co-Skewness Risk: Regression-Based Estimates for Portfolios Sorted on Co-Skewness

The table shows the results of Fama-MacBeth regressions using monthly returns. Each month, we estimate the CAPM and co-skewness betas jointly using a 60-month rolling window and calculate the value-weighted returns of 25 double-sorted COSK/CAPM portfolios. Using these portfolios as test assets, we run time-series regressions of the returns on the factors and then run cross-sectional regressions using returns for the next month to estimate the prices of covariance and co-skewness risks. Panel A reports the time-series average of the monthly estimates and the Fama-MacBeth t-statistics with Newey-West correction for serial correlation, using 1 lag. The $\lambda$s are multiplied by 100 for expositional convenience. Panel B presents the pre- and post-formation betas. The pre-formation betas correspond to the average across the entire sample of the value-weighted betas, and the averages of the monthly value-weighted t-statistics are given in parentheses. The post-formation betas and their t-statistics are obtained by regressing the monthly sorted portfolio returns on the factors. We consider two periods, 1986-2012 and 1966-2012.

Panel A: Co-Skewness Risk

| | 25 COSK/CAPM | | | |
| | 1986-2012 | | 1966-2012 | |
|---|---|---|---|---|
| $\lambda^0$ | 0.121 | -0.166 | 0.242 | 0.087 |
| | (0.42) | (-0.54) | (1.13) | (0.37) |
| $\lambda^{MKT}$ | 0.468 | 0.794 | 0.246 | 0.441 |
| | (1.18) | (1.99) | (0.83) | (1.47) |
| $\lambda^{HML}$ | | 0.183 | | 0.075 |
| | | (0.84) | | (0.47) |
| $\lambda^{SMB}$ | | 0.062 | | 0.087 |
| | | (0.33) | | (0.61) |
| $\lambda^{MOM}$ | | 0.893 | | 0.779 |
| | | (2.57) | | (3.13) |
| $\lambda^{COSK}$ | -0.014 | -0.034 | -0.028 | -0.030 |
| | (-0.44) | (-1.02) | (-1.25) | (-1.28) |
| | | | | |
| Adj $R^2$ | 22.62 | 30.89 | 25.57 | 32.43 |

Panel B: Pre- and Post-Formation Betas

| | Co-Skewness Factor | | | | CAPM Factor | | | |
| | 1986-2012 | | 1966-2012 | | 1986-2012 | | 1966-2012 | |
| Rank | Pre-Form. | Post-Form. | Pre-Form. | Post-Form. | Pre-Form. | Post-Form. | Pre-Form. | Post-Form. |
|---|---|---|---|---|---|---|---|---|
| 1 | -9.831 | 0.502 | -8.995 | 0.319 | 0.246 | 0.817 | 0.383 | 0.649 |
| | (-1.80) | (0.63) | (-1.78) | (0.68) | (1.43) | (12.59) | (2.34) | (22.86) |
| 2 | -4.289 | 0.331 | -3.883 | 0.226 | 0.617 | 1.045 | 0.736 | 0.861 |
| | (-1.18) | (0.80) | (-1.10) | (0.83) | (3.04) | (21.02) | (3.87) | (35.34) |
| 3 | -1.235 | 0.419 | -1.031 | 0.394 | 0.913 | 1.207 | 1.026 | 1.027 |
| | (-0.41) | (1.08) | (-0.34) | (1.44) | (4.16) | (25.21) | (4.79) | (41.42) |
| 4 | 1.618 | -0.011 | 1.663 | 0.079 | 1.254 | 1.396 | 1.351 | 1.229 |
| | (0.47) | (-0.12) | (0.51) | (0.23) | (5.15) | (32.75) | (5.51) | (47.15) |
| 5 | 6.368 | -0.626 | 6.175 | -0.434 | 1.893 | 1.700 | 1.932 | 1.566 |
| | (1.55) | (-1.32) | (1.63) | (-1.28) | (5.59) | (30.67) | (5.78) | (42.35) |

## Table 2: Regression-Based Estimates of the Price of Co-Skewness Risk

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

$$R_{j,t} - R_{f,t} = \alpha_{j,t} + \beta_{j,t}^{MKT} R_{m,t} + \beta_{j,t}^{HML} R_{HML,t} + \beta_{j,t}^{SMB} R_{SMB,t} + \beta_{j,t}^{MOM} R_{MOM,t} + \beta_{j,t}^{COSK} R_{m,t}^2 + \varepsilon_{j,t}.$$

We then run the following cross-sectional regression using the estimated betas and returns for the next month

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

We report the mean (in percentage) of the estimates and the Fama-MacBeth t-statistics with Newey-West correction for serial correlation, using 1 lag. The λs are multiplied by 100 for expositional convenience. We consider two periods, 1986-2012 and 1966-2012, and four sets of test assets.

Panel A: 1986 - 2012

|  | 25 Size/BM | | 25 size/Mom | | 25 size/STR | | 25 size/LTR | |
|---|---|---|---|---|---|---|---|---|
| $\lambda^0$ | 0.968 | 1.197 | 0.069 | 0.871 | 0.215 | 0.111 | 0.133 | 0.274 |
|  | (2.28) | (3.33) | (0.17) | (2.81) | (0.59) | (0.26) | (0.45) | (0.85) |
| $\lambda^{MKT}$ | -0.396 | -0.600 | 0.486 | -0.154 | 0.372 | 0.477 | 0.569 | 0.417 |
|  | (-0.85) | (-1.58) | (1.07) | (-0.42) | (0.83) | (1.05) | (1.67) | (1.14) |
| $\lambda^{HML}$ |  | 0.043 |  | 0.103 |  | -0.090 |  | 0.125 |
|  |  | (0.24) |  | (0.53) |  | (-0.50) |  | (0.60) |
| $\lambda^{SMB}$ |  | 0.274 |  | -0.177 |  | 0.116 |  | 0.101 |
|  |  | (1.45) |  | (-0.68) |  | (0.39) |  | (0.43) |
| $\lambda^{MOM}$ |  | 0.737 |  | 0.532 |  | -0.507 |  | 0.113 |
|  |  | (1.82) |  | (1.90) |  | (-1.18) |  | (0.35) |
| $\lambda^{COSK}$ | -0.080 | -0.059 | -0.148 | -0.037 | -0.008 | 0.023 | -0.058 | -0.035 |
|  | (-1.21) | (-1.16) | (-2.45) | (-0.87) | (-0.13) | (0.42) | (-0.92) | (-0.60) |
| Adj $R^2$ | 26.75 | 46.30 | 25.05 | 54.54 | 29.71 | 48.70 | 21.58 | 43.38 |

Panel B: 1966 - 2012

|  | 25 Size/BM | | 25 size/Mom | | 25 size/STR | | 25 size/LTR | |
|---|---|---|---|---|---|---|---|---|
| $\lambda^0$ | 0.755 | 0.892 | 0.107 | 0.807 | -0.538 | -0.152 | 0.252 | 0.647 |
|  | (2.40) | (3.48) | (0.35) | (3.42) | (-1.68) | (-0.49) | (1.11) | (2.67) |
| $\lambda^{MKT}$ | -0.251 | -0.440 | 0.397 | -0.258 | 1.017 | 0.555 | 0.364 | -0.123 |
|  | (-0.70) | (-1.60) | (1.17) | (-0.96) | (2.70) | (1.67) | (1.32) | (-0.44) |
| $\lambda^{HML}$ |  | 0.193 |  | 0.227 |  | 0.138 |  | 0.229 |
|  |  | (1.39) |  | (1.56) |  | (0.96) |  | (1.50) |
| $\lambda^{SMB}$ |  | 0.372 |  | -0.068 |  | 0.199 |  | 0.231 |
|  |  | (2.72) |  | (-0.39) |  | (0.92) |  | (1.42) |
| $\lambda^{MOM}$ |  | 0.433 |  | 0.680 |  | -1.116 |  | 0.042 |
|  |  | (1.58) |  | (3.56) |  | (-3.71) |  | (0.17) |
| $\lambda^{COSK}$ | -0.046 | -0.045 | -0.146 | -0.060 | -0.027 | 0.034 | -0.040 | -0.049 |
|  | (-1.05) | (-1.32) | (-3.56) | (-2.12) | (-0.58) | (0.91) | (-0.95) | (-1.35) |
| Adj $R^2$ | 30.16 | 48.95 | 26.37 | 54.40 | 30.26 | 49.61 | 23.07 | 43.47 |

## Table 3: The Option-Based Price of Co-Skewness Risk

The table provides descriptive statistics for the physical and risk-neutral expectations and the price of co-skewness risk. The data are monthly. The physical second moment is estimated using a HAR model and the risk-neutral second moment is proxied by the VIX squared. The time-varying price of co-skewness risk is equal to the spread between the physical and risk-neutral moments. The moments and prices of risk are multiplied by 100 for expositional convenience. The data are monthly and the sample period is from January 1986 to December 2012.

|  | $E_t^P[R_{m,t+1}^2]$ | $E_t^Q[R_{m,t+1}^2]$ | $E_t^P[R_{m,t+1}^2] - E_t^Q[R_{m,t+1}^2]$ |
|---|---|---|---|
| mean | 0.1675 | 0.4381 | -0.2707 |
| std | 0.2158 | 0.4133 | 0.3105 |
| skew | 10.5092 | 3.4318 | -3.5761 |
| kurt | 145.8792 | 19.0463 | 23.1180 |
| $\rho(1)$ | 0.4013 | 0.7525 | 0.4949 |

## Table 4: Out-of-Sample Fit

We compare the out-of-sample performance of several competing models. The price of risk is estimated using cross-sectional regressions or historical risk premia. Using historical risk premia, we provide out-of-sample predictions for the $CAPM$, the model with a co-skewness premium $COSK$, and the model with market and co-skewness factors $CAPM + COSK$. Using cross-sectional regressions, we provide predictions for the $CAPM$, the model with market and co-skewness factors $CAPM + COSK$, and the Fama-French three-factor model $FF$. We also use a hybrid model with regression-based co-variance premium and historical co-skewness premium. This model is referred to as $CSCAPM + COSK$. We consider four sets of portfolios. We compute out-of-sample $R$-squares for each portfolio and report the average. The sample period is from January 1986 to December 2012.

Out of Sample R-squares

| Prices of Risk Estimated from Historical Risk Premia | | | | |
|---|---|---|---|---|
| | CAPM | COSK | CAPM +COSK | CSCAPM + COSK |
| 25 Size/Book-to-Market | -0.252 | 1.690 | -0.005 | -2.061 |
| 25 Size/Momentum | -0.553 | 0.377 | -1.623 | 0.353 |
| 25 Size/Short-Term Reversal | -0.301 | 1.362 | -0.676 | 0.304 |
| 25 Size/Long-Term Reversal | -0.420 | 0.888 | -0.533 | 0.720 |

| Prices of Risk Estimated from Cross-Sectional Regressions | | | |
|---|---|---|---|
| | CAPM | CAPM +COSK | FF |
| 25 Size/Book-to-Market | -3.905 | -4.534 | -7.141 |
| 25 Size/Momentum | 0.879 | -0.214 | -6.237 |
| 25 Size/Short-Term Reversal | 0.469 | 0.524 | -0.691 |
| 25 Size/Long-Term Reversal | 0.440 | 0.509 | -0.750 |

### Table 5: Out-of-Sample Fit During Good and Bad Times

We compare the out-of-sample performance of several competing models in good and bad times. Bad (Good) times correspond to years with negative (positive) market returns. The price of risk is estimated using cross-sectional regressions or historical risk premia. As in Table 4, we consider seven models and four sets of portfolios. We compute out-of-sample R-squares for each portfolio and report the average. The sample period is from January 1986 to December 2012.

Panel A: Out of Sample R-squares During Bad Times

Prices of Risk Estimated from Historical Risk Premia

|  | CAPM | COSK | CAPM +COSK | CSCAPM + COSK |
|---|---|---|---|---|
| 25 Size/Book-to-Market | 1.274 | 4.809 | 1.502 | 5.758 |
| 25 Size/Momentum | 0.527 | 3.375 | -0.577 | 1.363 |
| 25 Size/Short-Term Reversal | 0.237 | 4.066 | -0.295 | 1.976 |
| 25 Size/Long-Term Reversal | 1.397 | 4.827 | 1.770 | 1.902 |

Prices of Risk Estimated from Cross-Sectional Regressions

|  | CAPM | CAPM +COSK | FF |
|---|---|---|---|
| 25 Size/Book-to-Market | 4.610 | 4.806 | -0.556 |
| 25 Size/Momentum | 2.372 | 0.589 | 3.816 |
| 25 Size/Short-Term Reversal | 2.395 | 2.745 | 0.099 |
| 25 Size/Long-Term Reversal | 1.346 | 1.915 | 1.951 |

Panel B: Out of Sample R-squares During Good Times

Prices of Risk Estimated from Historical Risk Premia

|  | CAPM | COSK | CAPM +COSK | CSCAPM + COSK |
|---|---|---|---|---|
| 25 Size/Book-to-Market | -1.275 | -0.545 | -0.982 | -7.816 |
| 25 Size/Momentum | -1.268 | -1.661 | -2.342 | -0.271 |
| 25 Size/Short-Term Reversal | -0.722 | -0.693 | -0.970 | -0.931 |
| 25 Size/Long-Term Reversal | -1.666 | -1.702 | -1.953 | 0.039 |

Prices of Risk Estimated from Cross-Sectional Regressions

|  | CAPM | CAPM +COSK | FF |
|---|---|---|---|
| 25 Size/Book-to-Market | -10.243 | -11.468 | -12.289 |
| 25 Size/Momentum | -0.078 | -0.681 | -13.351 |
| 25 Size/Short-Term Reversal | -0.970 | -1.122 | -1.264 |
| 25 Size/Long-Term Reversal | -0.180 | -0.438 | -2.673 |

## Table 6: The Price of Co-Skewness Risk and Out-of-Sample Performance: Robustness

In Panel A, we provide estimates of the price of co-skewness risk using alternative estimators of the physical and risk-neutral second moments. In Panel B, we document the out-of-sample performance of the model with co-skewness risk ($COSK$) using these different moment estimators. In Panel B we consider four sets of portfolios. The moments and prices of risk are multiplied by 100 for expositional convenience. The sample periods differ dependent on data availability.

Panel A: Price of Co-Skewness Risk

|  | Sample Period | $E_t^P[R_{m,t+1}^2]$ | $E_t^Q[R_{m,t+1}^2]$ | $E_t^P[R_{m,t+1}^2] - E_t^Q[R_{m,t+1}^2]$ |
|---|---|---|---|---|
| NGARCH + VIX | 1986 - 2012 | 0.2960 | 0.4381 | -0.1421 |
| NGARCH + BKM | 1996 - 2012 | 0.3445 | 0.5153 | -0.1708 |
| Heston + VIX | 1986 - 2012 | 0.1539 | 0.4386 | -0.2847 |
| Heston + BKM | 1996 - 2012 | 0.2132 | 0.5167 | -0.3035 |
| AR + VIX | 1986 - 2012 | 0.3152 | 0.4381 | -0.1229 |
| AR + BKM | 1996 - 2012 | 0.3462 | 0.5153 | -0.1691 |
| HAR + BKM | 1996 - 2012 | 0.1990 | 0.5153 | -0.3163 |

Panel B: Out-of-Sample R-squares

|  | Sample Period | 25 Size/BM | 25 Size/Mom | 25 Size/STR | 25 Size/LTR |
|---|---|---|---|---|---|
| NGARCH + VIX | 1986 - 2012 | 1.976 | 0.762 | 1.679 | 1.236 |
| NGARCH + BKM | 1996 - 2012 | 2.165 | 0.345 | 1.709 | 1.056 |
| Heston + VIX | 1986 - 2012 | 1.521 | 0.492 | 1.243 | 0.750 |
| Heston + BKM | 1996 - 2012 | 1.437 | -0.070 | 1.076 | 0.120 |
| AR + VIX | 1986 - 2012 | 0.974 | 0.618 | 1.352 | 0.333 |
| AR + BKM | 1996 - 2012 | 0.925 | 0.278 | 1.442 | -0.091 |
| HAR + BKM | 1996 - 2012 | 1.999 | 0.203 | 1.602 | 0.765 |

**Table 7: Option-Based Prices of Co-Skewness and Co-Kurtosis Risk. Out-of-Sample Fit.**

We compare the out-of-sample performance of several competing models. The price of risk is estimated using cross-sectional regressions or historical risk premia. Using historical risk premia, we provide out-of-sample predictions for the $CAPM$, the model with a co-skewness premium $COSK$, the model with a co-kurtosis premium $COKU$, and the model with market and co-skewness factors and market and co-kurtosis factors, $CAPM + COSK$ and $CAPM + COKU$ respectively. Using cross-sectional regressions, we provide predictions for the $CAPM$, $CAPM + COSK$, $CAPM + COKU$, and the Fama-French three-factor model $FF$. We also use hybrid models with regression-based co-variance premium and historical co-skewness and co-kurtosis premiums, $CSCAPM + COSK$ and $CSCAPM + COKU$. We use the Bakshi-Madan (2000) method and zero physical skew to obtain the time-series of the price of cokurtosis risk used in the $COKU$, $CAPM + COKU$, and $CSCAPM + COKU$ specifications. We use the HAR+VIX benchmark model in order to obtain the time-series of the price of co-skewness risk used in the $COSK$, $CAPM + COSK$, and $CSCAPM + COSK$ specifications. We consider four sets of portfolios and report out-of-sample $R$-squares. The sample period is from January 1996 to December 2012.

Out-of-Sample R-squares

Prices of Risk Estimated from Historical Risk Premia

|  | CAPM | COSK | COKU | CAPM +COSK | CAPM +COKU | CSCAPM +COSK | CSCAPM +COKU |
|---|---|---|---|---|---|---|---|
| 25 Size/Book-to-Market | -0.402 | 2.279 | 0.934 | 0.100 | -0.339 | -0.011 | -2.674 |
| 25 Size/Momentum | -0.732 | 0.630 | 0.400 | -1.855 | -1.080 | 1.035 | 0.763 |
| 25 Size/Short-Term Reversal | -0.418 | 1.847 | 1.007 | -0.653 | -0.660 | 0.811 | 0.645 |
| 25 Size/Long-Term Reversal | -0.645 | 1.143 | 0.214 | -0.771 | -0.977 | 1.579 | 0.685 |

Prices of Risk Estimated from Cross-Sectional Regressions

|  | CAPM | CAPM +COSK | CAPM +COKU | FF |
|---|---|---|---|---|
| 25 Size/Book-to-Market | -1.944 | -2.472 | -3.931 | -7.967 |
| 25 Size/Momentum | 1.596 | 0.961 | 0.615 | -4.681 |
| 25 Size/Short-Term Reversal | 1.048 | 1.736 | 0.008 | -0.393 |
| 25 Size/Long-Term Reversal | 0.998 | 1.046 | 0.179 | 0.055 |

## Table 8: The Price of Co-Kurtosis Risk and Out-of-Sample Performance: Robustness

Panel A provides estimates of the co-kurtosis price of risk using the $COKU$ model and alternative measures of the physical third moment. The estimates of the third moment and the price of co-kurtosis risk reported in Panel A are multiplied by 100 for expositional convenience. Panel B reports out-of-sample $R$-squares using these alternative estimates. Panel B considers four sets of portfolios. We compute out-of-sample $R$-squares for each portfolio and report the average. The sample period is from January 1996 to December 2012.

Panel A: The Price of Co-Kurtosis Risk

|  | $E_t^P[R_{m,t+1}^3]$ | $E_t^Q[R_{m,t+1}^3]$ | $E_t^P[R_{m,t+1}^3] - E_t^Q[R_{m,t+1}^3]$ |
|---|---|---|---|
| const skew (daily) | -0.0006 | -0.0220 | 0.0214 |
| const skew (monthly) | -0.0015 | -0.0220 | 0.0205 |
| Jondeau and Rockinger (2003) | -0.0104 | -0.0220 | 0.0116 |

Panel B: Out-of-Sample R-squares

|  | 25 Size/BM | 25 Size/Mom | 25 Size/STR | 25 Size/LTR |
|---|---|---|---|---|
| const skew (daily) | 1.088 | 0.510 | 1.098 | 0.360 |
| const skew (monthly) | 1.038 | 0.385 | 0.974 | 0.172 |
| Jondeau and Rockinger (2003) | 0.396 | -0.432 | 0.397 | -0.332 |

# Research Papers
# 2013



| 2015-36: | Markku Lanne and Henri Nyberg: Nonlinear dynamic interrelationships between real activity and stock returns |
|---|---|
| 2015-37: | Markku Lanne and Jani Luoto: Estimation of DSGE Models under Diffuse Priors and Data-Driven Identification Constraints |
| 2015-38: | Lorenzo Boldrini and Eric Hillebrand: Supervision in Factor Models Using a Large Number of Predictors |
| 2015-39: | Lorenzo Boldrini and Eric Hillebrand: The Forecasting Power of the Yield Curve, a Supervised Factor Model Approach |
| 2015-40: | Lorenzo Boldrini: Forecasting the Global Mean Sea Level, a Continuous-Time State-Space Approach |
| 2015:41: | Yunus Emre Ergemen and Abderrahim Taamouti: Parametric Portfolio Policies with Common Volatility Dynamics |
| 2015-42: | Mikkel Bennedsen: Rough electricity: a new fractal multi-factor model of electricity spot prices |
| 2015-43: | Mikkel Bennedsen, Asger Lunde and Mikko S. Pakkanen: Hybrid scheme for Brownian semistationary processes |
| 2015-44: | Jonas Nygaard Eriksen: Expected Business Conditions and Bond Risk Premia |
| 2015-45: | Kim Christensen, Mark Podolskij, Nopporn Thamrongrat and Bezirgen Veliyev: Inference from high-frequency data: A subsampling approach |
| 2015-46: | Asger Lunde, Anne Floor Brix and Wei Wei: A Generalized Schwartz Model for Energy Spot Prices - Estimation using a Particle MCMC Method |
| 2015-47: | Annastiina Silvennoinen and Timo Teräsvirta: Testing constancy of unconditional variance in volatility models by misspecification and specification tests |
| 2015-48: | Harri Pönkä: The Role of Credit in Predicting US Recessions |
| 2015-49: | Palle Sørensen: Credit policies before and during the financial crisis |
| 2015-50: | Shin Kanaya: Uniform Convergence Rates of Kernel-Based Nonparametric Estimators for Continuous Time Diffusion Processes: A Damping Function Approach |
| 2015-51: | Tommaso Proietti: Exponential Smoothing, Long Memory and Volatility Prediction |
| 2015-52: | Mark Podolskij, Christian Schmidt and Mathias Vetter: On U- and V-statistics for discontinuous Itô semimartingale |
| 2015-53: | Mark Podolskij and Nopporn Thamrongrat: A weak limit theorem for numerical approximation of Brownian semi-stationary processes |
| 2015-54: | Peter Christoffersen, Mathieu Fournier, Kris Jacobs and Mehdi Karoui: Option-Based Estimation of the Price of Co-Skewness and Co-Kurtosis Risk |