JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 56, No. 1, Feb. 2021, pp. 65–91

doi:10.1017/S002210902000023X

# Option-Based Estimation of the Price of Coskewness and Cokurtosis Risk

Peter Christoffersen
*(deceased)*

Mathieu Fournier
*HEC Montréal*
mathieu.fournier@hec.ca

Kris Jacobs
*University of Houston*
kjacobs@bauer.uh.edu (corresponding author)

Mehdi Karoui
*Ontario Municipal Employees Retirement System (OMERS)*
mehdi.karoui@mail.mcgill.ca

## Abstract

We show that the prices of risk for factors that are nonlinear in the market return can be obtained using index option prices. The price of coskewness risk corresponds to the market variance risk premium, and the price of cokurtosis risk corresponds to the market skewness risk premium. Option-based estimates of the prices of risk lead to reasonable values of the associated risk premia. An analysis of factor models with coskewness risk indicates that the new estimates of the price of risk improve the models' performance compared with regression-based estimates.

## I. Introduction

Kraus and Litzenberger (1976) show that if investors care about portfolio skewness, coskewness is relevant for pricing assets, in addition to covariation with the market portfolio. If investors care about portfolio kurtosis, cokurtosis is also relevant.[1] We propose a strategy to estimate the price of coskewness and cokurtosis risk that avoids the problems inherent in the use of 2-stage cross-sectional or Fama–MacBeth regressions. Our approach can also be used to estimate the price of other risks, provided that they are nonlinear functions of the market return, and it can accommodate the presence of other pricing factors.

Peter Christoffersen, our dear friend and coauthor, tragically passed away prior to the publication of this article. We thank the Bank of Canada, the Global Risk Institute, and the Social Sciences and Humanities Research Council (SSHRC) for financial support. For helpful comments, we thank Tim Bollerslev, Robert Dittmar (the referee), Adlai Fisher, Bryan Kelly, Eric Renault, Michael Rockinger, Sang Byung Seo, Bas Werker, Xiaoyan Zhang, and seminar participants at the 2015 American Finance Association (AFA) Meeting and the 2014 Society for Financial Econometrics (SoFiE) meeting.

[1]For important contributions to the literature on coskewness and cokurtosis, see, for example, Ang, Chen, and Xing (2006), Bansal and Viswanathan (1993), Dittmar (2002), Guidolin and Timmermann (2008), Harvey and Siddique (2000), Leland (1997), Lim (1989), Schneider (2015), and Scott and Horvath (1980). See also Rubinstein (1973) and Golec and Tamarkin (1998) for related work.

Our approach explicitly uses restrictions on the pricing of contingent claims as well as stocks. This allows us to derive explicit expressions for the time-varying price of risk of the exposure to any nonlinear function of the market return. The price of coskewness risk corresponds to the spread between the physical and the risk-neutral second moment, the market variance risk premium. Similarly, the price of cokurtosis risk is given by the spread between the physical and the risk-neutral third moments, the market skewness risk premium.

The theoretical insights that motivate this approach are available in the existing literature. Consider the well-known special case where the stochastic discount factor (SDF) is a linear function of the market return, which corresponds to the capital asset pricing model (CAPM). In this case, our approach implies that the price of covariance risk corresponds to the spread between the physical and risk-neutral first moments. This equals the market return minus the risk-free rate, which is, of course, the classic CAPM result. Although information from index options is not useful for this special case with a linear SDF, we show that for all other cases when the SDF is nonlinear, 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 because index options are among the most heavily traded contracts on the market. This makes our theoretical results very practical to implement.

Despite the obvious intuitive appeal of coskewness and cokurtosis, there seems to be no widespread consensus on their empirical relevance for cross-sectional asset pricing. We address one of the most important drawbacks of the approach, which is the measurement of the risk premia. Measurement is especially difficult when analyzing *conditional* coskewness and cokurtosis.[2] Most existing articles estimate and test the importance of coskewness and cokurtosis using 2-stage cross-sectional or Fama–MacBeth regressions. This approach necessitates the estimation of coskewness in a first stage. These estimates are subsequently used in the second-stage cross-sectional regression. It is well known that the estimation of coskewness in the first-stage regression is noisy, and these errors carry over in the second-stage cross-sectional regression.[3] Although these problems apply to virtually all implementations of cross-sectional models, including the CAPM, they may be especially serious in the case of coskewness and cokurtosis. The higher the moment, the more difficult it is to precisely estimate loadings, which leads to biases in the cross-sectional estimates of the price of risk that are potentially much larger than in the case of the CAPM or the Fama–French 3-factor model.[4]

We empirically investigate the performance of our approach for the pricing of coskewness and cokurtosis risk. Using monthly data for the periods 1986–2012 and 1996–2012, respectively, we find that our conditional prices of coskewness and

[2]Kraus and Litzenberger (1976) provide an unconditional empirical analysis of coskewness. For a classic example of a conditional analysis, see, for instance, Harvey and Siddique (2000). For a general comparison of the performance of conditional and unconditional approaches, see Ghysels (1998).

[3]See, for example, Jagannathan and Wang (1998), Shanken (1992), Kan and Zhang (1999), Kleibergen (2009), and Gospodinov, Kan, and Robotti (2014).

[4]Note that our empirical implementation also relies on the first-stage loadings to estimate the factor exposures, but crucially, we do not use these (noisy) loadings to estimate the price of risk in a second stage. Instead, we obtain the price of risk using option-implied information, leading to a truly conditional and much less noisy estimate. We leave the improved estimation of loadings for future research.

cokurtosis risk are stable and have the expected sign in almost every month in our sample. On average, both estimated prices of risk are larger in absolute value than the traditional estimates obtained using a 2-stage Fama–MacBeth approach. Although the average prices of risk obtained using the Fama–MacBeth approach most often have the theoretically anticipated signs, the estimates are noisy, and reliable conditional estimates are not available. We evaluate the cross-sectional performance of our newly proposed estimates out-of-sample, and we find that they outperform regression-based estimates of the price of risk.

In addition to the existing work on coskewness and cokurtosis discussed previously, our work is also related to several other strands of literature. Several articles investigate the importance of conditional skewness for the cross section of stock returns and asset allocation. Recent work includes Conrad, Dittmar, and Ghysels (2013), Ghysels, Plazzi, and Valkanov (2016), and Schneider, Wagner, and Zechner (2020). Another literature uses higher moments of index returns extracted from option data as pricing factors. This literature follows Ang, Hodrick, Xing, and Zhang (2006), who show that the Volatility Index (VIX) is a priced factor in the cross section of stock returns. Chang, Christoffersen, and Jacobs (2013) show that option-implied skewness is priced in the cross section of stock returns, and this finding can be explained by our results on cokurtosis. Cremers, Halling, and Weinbaum (2015) show that aggregate jump risk and aggregate volatility risk are priced in the cross section of stock returns. Martin (2017) shows that a volatility index that can be computed using option prices provides a lower bound on the equity premium. Chabi-Yo (2018) investigates the cross section of stock returns using option-implied moments of the SDF.

Our results also provide a theoretical motivation for the role of the variance risk premium in asset pricing. Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Marrone, Xu, and Zhou (2014) demonstrate that the variance risk premium is one of the best predictors of market returns.[5] Other articles investigate the empirical relation between the variance risk premium and the cross section of stock returns. Han and Zhou (2011) document return differentials for stocks with different variance risk premiums. Their focus is completely different because they investigate the firm-level variance risk premium rather than the role of the market variance risk premium in factor models. Bali and Zhou (2016) investigate the performance of market variance risk in the cross section of stocks. Starting from a consumption-based model with Epstein–Zin–Weil utility, they show that the market variance risk premium is a risk factor and interpret it as a proxy for economic uncertainty. Bali and Zhou (2016) also theoretically demonstrate the relevance of the market variance risk premium for the cross section of stock returns. However, in their approach, the market variance risk premium is a pricing factor, whereas in our setup, it is the price of risk.

Our pricing results critically depend on the functional form of the SDF as a function of market returns. Several recent articles empirically investigate whether the SDF is monotonic in market returns. Bakshi, Madan, and Panayotov (2010), Patton and Timmermann (2010), and Chaudhuri and Schroder (2015) reject a monotonic relationship and provide strong evidence that the pricing kernel is U-shaped. Nonlinear SDFs can be motivated by the existence of investors'

[5]Bollerslev et al. (2009) also motivate these empirical results with a general equilibrium model.

preference for higher-order moments. For instance, a quadratic SDF is obtained in equilibrium when the third derivative of the marginal agent's utility function is nonzero, which is closely related to skewness preferences. Similarly, a nonzero fourth derivative of the utility function is related to kurtosis preferences.

Finally, our work is also related to the recent literature on disaster risk and heavy-tailed shocks in equilibrium models. Extending the approaches of Rietz (1988) and Barro (2006), Gabaix (2012), Gourio (2012), and Wachter (2013) argue that such heavy-tailed shocks to endowments or productivity can help explain long-standing puzzles in asset pricing.[6] Wachter (2013) notes that these articles are related to the coskewness and cokurtosis literature that examines the effects of nonnormalities on risk premia and the cross section of returns. Bollerslev and Todorov (2011) find that the largest part of the variance risk premium is due to jump and crash risk. Drechsler and Yaron (2011) study the impact of heavy-tailed shocks in a long-run risk model and emphasize the importance of the variance risk premium in capturing attitudes toward uncertainty. They explicitly discuss the importance of combining derivatives information and information on the underlying to capture empirically relevant risk measures. Our article theoretically justifies this approach and demonstrates its relevance for the cross section of returns.

The article proceeds as follows: Section II discusses our alternative approach as applied to the measurement of the price of coskewness risk. Section III presents empirical results for coskewness risk and compares the pricing performance of the new approach with regression-based estimates. Section IV discusses the predictive performance of the new estimates. Section V applies the new approach to the measurement of cokurtosis risk as well as more general nonlinear market risk. Section VI concludes. The Supplementary Material collects additional empirical results and robustness tests.

## II. Measuring Coskewness Risk: An Option-Based Approach

In this section, we investigate an asset pricing model in which cross-sectional differences in expected returns between assets are determined by their exposure to the squared market return in addition to the market return itself. We proceed to propose an option-based approach to measuring the price of risk for the nonlinear (coskewness) exposure.

### A. A Quadratic SDF

The absence of arbitrage implies the existence of an SDF, $m_{t+1}$, that prices any asset with a risky return, $R_{j,t+1}$, using the moment condition

$$E_t^P\left[\left(1+R_{j,t+1}\right)m_{t+1}\right]=1, \tag{1}$$

[6]For additional work on equilibrium models of disaster risk, see Chen, Joslin, and Tran (2012), Julliard and Ghosh (2012), and Liu, Pan, and Wang (2005). For recent work on tail risk, see Bollerslev and Todorov (2011) and Kelly and Jiang (2014).

where $E_t^P(\cdot)$ denotes the expectation under the physical probability measure. We assume that the SDF can be written as a representative investor's marginal rate of substitution between current and future wealth,

$$m_{t+1} = \frac{U'(W_{t+1})}{U'(W_t)}, \tag{2}$$

where $U'(\cdot)$ is marginal utility, and $W$ is aggregate wealth. We do not make a specific assumption regarding the form of the representative investor's utility function; instead, we take a second-order Taylor approximation for $U'(W_{t+1})$ around $W_t$ to write[7]

$$m_{t+1} \approx 1 + \left(\frac{U''(W_t)W_t}{U'(W_t)}\right) R_{m,t+1} + \left(\frac{U'''(W_t)W_t^2}{2U'(W_t)}\right) R_{m,t+1}^2, \tag{3}$$

where we have used $W_{t+1} = W_t(1+R_{m,t+1})$, and $R_{m,t+1}$ denotes the stock market return, which we will use as a proxy for the return on the wealth portfolio.[8] From equation (1), we have $E_t^P(m_{t+1}) = 1/\left(1+R_{f,t}\right)$ for the risk-free rate, which gives

$$\frac{1}{\left(1+R_{f,t}\right)} \approx 1 + \left(\frac{U''(W_t)W_t}{U'(W_t)}\right) E_t^P(R_{m,t+1}) + \left(\frac{U'''(W_t)W_t^2}{2U'(W_t)}\right) E_t^P\left(R_{m,t+1}^2\right). \tag{4}$$

Combining equations (3) and (4), we obtain the following form for the SDF:

$$m_{t+1} \approx a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P\left(R_{m,t+1}^2\right)\right), \tag{5}$$

where $a_t = 1/\left(1+R_{f,t}\right)$, $b_{1,t} = (U''(W_t)W_t)/U'(W_t)$, and $b_{2,t} = \left(U'''(W_t)W_t^2\right)/(2U'(W_t))$. Similar to Harvey and Siddique (HS) (2000), our setup is based on the assumption of $U'''(\cdot) \neq 0$, which implies a quadratic SDF. Dittmar (2002) assumes declining absolute risk aversion, which implies that $U'''(\cdot) > 0$. The performance of quadratic pricing kernels is studied further by Bansal and Viswanathan (1993) and Chabi-Yo (2008), among others. Chabi-Yo, Leisen, and Renault (2014) study a general equilibrium setting for skewness risk. They find that the quadratic SDF in equation (5) obtains as a special case in a single-period setting with a representative investor.[9]

## B. The Price of Coskewness Risk

Given the SDF in equation (5), 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

[7] See also, among others, Kraus and Litzenberger (1976), Barone-Adesi (1985), Harvey and Siddique (2000), and Dittmar (2002).

[8] We use this assumption because our objective is the use of index options to estimate risk premia. Bali and Zhou (2016) start from a consumption-based SDF and use Campbell's (1993), (1996) approach to substitute out consumption and obtain moments of the market return as priced risk factors.

[9] Allowing for heterogenous agents and/or multiperiod models results in more general prices of risk than the ones derived in our complete-market setup. Investor heterogeneity is studied in, among others, Constantinides and Duffie (1996) and Mitton and Vorkink (2007).

asset pricing literature focuses exclusively on the underlying assets. Our approach enables the specification of new estimators for the price of coskewness risk, harvesting the rich information in index option prices.

The existing literature contains several measures of coskewness risk, which all capture covariation between the stock return and the squared market return. Kraus and Litzenberger (KL) (1976) define coskewness risk by

$$E^P\left[\left(R_j-\overline{R}_j\right)\left(R_m-\overline{R}_m\right)^2\right]\Big/E^P\left[\left(R_m-\overline{R}_m\right)^3\right].$$

HS (2000) mainly focus on $\operatorname{cov}\left(R_j,R_m^2\right)$ in their theoretical analysis but consider four coskewness measures in their empirical analysis. Our measure of coskewness risk is the loading 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, which presents the pricing implications of the SDF defined in equation (5).

*Proposition 1.* In the absence of arbitrage opportunities, if the SDF has the following form:

$$(6)\qquad m_{t+1}=a_t+b_{1,t}\left(R_{m,t+1}-E_t^P(R_{m,t+1})\right)+b_{2,t}\left(R_{m,t+1}^2-E_t^P\left(R_{m,t+1}^2\right)\right),$$

then the cross-sectional pricing restriction on stock returns is

$$(7)\qquad E_t^P\left(R_{j,t+1}\right)-R_{f,t}=\lambda_t^{\text{MKT}}\beta_{j,t}^{\text{MKT}}+\lambda_t^{\text{COSK}}\beta_{j,t}^{\text{COSK}},$$

where $\beta_{j,t}^{\text{MKT}}$and $\beta_{j,t}^{\text{COSK}}$ 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, $\lambda_t^{\text{MKT}}$, is

$$(8)\qquad \lambda_t^{\text{MKT}}=E_t^P(R_{m,t+1})-R_{f,t},$$

and the price of coskewness risk, $\lambda_t^{\text{COSK}}$, is

$$(9)\qquad \lambda_t^{\text{COSK}}=E_t^P\left(R_{m,t+1}^2\right)-E_t^Q\left(R_{m,t+1}^2\right),$$

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

*Proof.* Linear factor models, in which the SDF 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 0 risk factors $\mathbf{f}$ satisfying

$$(10)\qquad E_t^P\left(R_{j,t+1}\right)-R_{f,t}=\lambda_t'\beta_{j,t},$$

where $\lambda_t'=(-1/a_t)\mathbf{b}_t'E_t^P\left(\mathbf{f}_{t+1}\mathbf{f}_{t+1}'\right)$, $\left(1+R_{f,t}\right)=(1/a_t)=\left(1/E_t^P(m_{t+1})\right)$, and $\beta_{j,t}=\left[E_t^P\left(\mathbf{f}_{t+1}\mathbf{f}_{t+1}'\right)\right]^{-1}E_t^P\left(\mathbf{f}_{t+1}R_{j,t+1}\right)$ (see, e.g., Cochrane (2005), p. 108). Because the pricing kernel prices all the assets, equation (10) 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 (10) to $\Psi$ gives

$$E_t^P(R_{\Psi,t+1}) - R_{f,t} = E_t^P\left(\frac{\Psi_{t+1}-P_t}{P_t}\right) - R_{f,t} = \lambda_t'\beta_{\Psi,t}, \tag{11}$$

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 $\beta_{\Psi,t}$, we have

$$E_t^P\left(\frac{\Psi_{t+1}-P_t}{P_t}\right) - R_{f,t} = \lambda_t'\left[E_t^P\left(\mathbf{f}_{t+1}\mathbf{f}_{t+1}'\right)\right]^{-1} E_t^P\left(\mathbf{f}_{t+1}\left(\frac{\Psi_{t+1}-P_t}{P_t}\right)\right). \tag{12}$$

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

$$\begin{aligned} E_t^P(\Psi_{t+1}) - P_t\left(1+R_{f,t}\right) &= \lambda_t'\left[E_t^P\left(\mathbf{f}_{t+1}\mathbf{f}_{t+1}'\right)\right]^{-1} E_t^P(\mathbf{f}_{t+1}\Psi_{t+1}) \\ &= \lambda_t'\widetilde{\beta}_{\Psi,t}, \end{aligned} \tag{13}$$

where $\widetilde{\beta}_{\Psi,t}$ 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 $P_t = E_t^Q(\Psi_{t+1})/\left(1+R_{f,t}\right)$. Therefore, we get

$$E_t^P(\Psi_{t+1}) - E_t^Q(\Psi_{t+1}) = \lambda_t'\widetilde{\beta}_{\Psi,t}. \tag{14}$$

To obtain the result in equation (8) from equation (14), 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 1 onto itself and 0 onto the other factors. Given the SDF in equation (6), this gives $\widetilde{\beta}_{\Psi,t} = [1\ 0]'$, and equation (14) reduces to equation (8). Similarly, using $\Psi_{t+1} \equiv R_{m,t+1}^2$, we obtain $\widetilde{\beta}_{\Psi,t} = [0\ 1]'$, which, applied to equation (14), gives the result in equation (9). □

Proposition 1 shows that the price of coskewness risk corresponds to the spread between the physical and risk-neutral second moments for the market return. A number of existing articles 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 coskewness risk.

Proposition 1 allows for separate identification of the price of covariance ($\lambda_t^{\mathrm{MKT}}$) and coskewness ($\lambda_t^{\mathrm{COSK}}$) 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 2-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 spread between the physical and risk-neutral market variances is often termed the *variance risk premium*. Proposition 1 suggests that the variance risk premium provides information about the price of coskewness risk. This insight is consistent with the finding of Bollerslev and Todorov (2011) that the largest part of the variance risk premium is due to jump and crash risk. The existing literature also shows that the variance risk premium is one of the best predictors of market returns (see, e.g., Bollerslev et al. (2009), (2014)). In our framework, the asset loadings $\beta_{j,t}^{\text{MKT}}$ and $\beta_{j,t}^{\text{COSK}}$ in equation (7) are obtained from a bivariate regression of the asset on the factors. Applying this to the market return means that the market return loads upon itself; our approach is therefore silent about the relation between the expected market return and other risk premiums, including the variance risk premium.

The existing empirical evidence clearly indicates that risk-neutral variance is larger than physical variance, therefore suggesting a negative price of coskewness risk.[10] A negative price of risk is consistent with theory. Assets with lower (more negative) coskewness decrease the total skewness of the portfolio and increase the likelihood of extreme losses. Assets with lower coskewness 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). We follow Bakshi and Madan and Bakshi et al., who use log returns. Another implementation issue is that throughout our empirical work, we use both centered and uncentered moments, and the results are very similar.

Finally, although our approach to estimating the price of coskewness risk is different from the existing literature and the loadings are defined (and/or scaled) differently, the implications for the risk premia on the assets are, of course, the same. Using the fact that $E_t^P(R_{m,t+1}) - R_{f,t} = \lambda_t^{\text{MKT}}$ and $E_t^P\left(R_{m,t+1}^2\right) - E_t^Q\left(R_{m,t+1}^2\right) = \lambda_t^{\text{COSK}}$, we can rewrite equation (7) of Proposition 1 as follows:

$$E_t^P\left(R_{j,t+1}\right) - R_{f,t} = \beta_{j,t}^{\text{MKT}}\left[E_t^P(R_{m,t+1}) - R_{f,t}\right] + \beta_{j,t}^{\text{COSK}}\left[E_t^P\left(R_{m,t+1}^2\right) - E_t^Q\left(R_{m,t+1}^2\right)\right], \tag{15}$$

which can also be written as

$$E_t^P\left(R_{j,t+1}\right) - R_{f,t} = c_{j,t} + \beta_{j,t}^{\text{MKT}}\left[E_t^P(R_{m,t+1}) - R_{f,t}\right] + \beta_{j,t}^{\text{COSK}} E_t^P\left[\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right)^2\right], \tag{16}$$

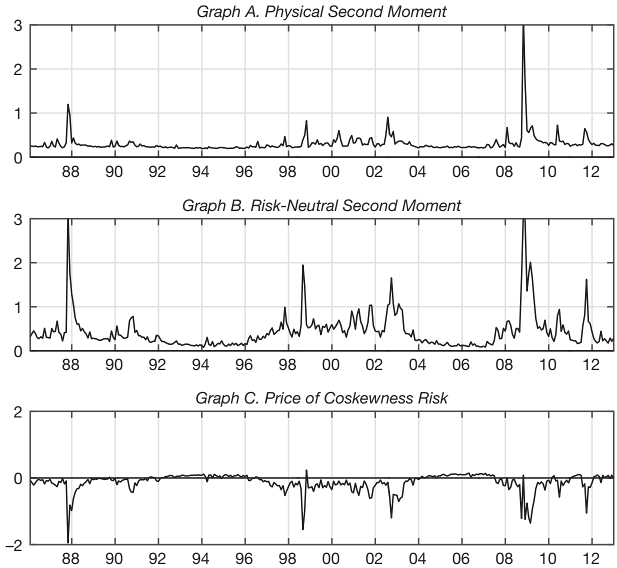
where $c_{j,t} = \beta_{j,t}^{\text{COSK}}\left[\left(E_t^P(R_{m,t+1})\right)^2 - E_t^Q\left(R_{m,t+1}^2\right)\right]$. Equation (16) directly establishes the link between our approach and existing work. For instance, it is equivalent to equation (6) in KL (1976), and it can be rewritten as equation (8) in HS (2000).

[10] See, for instance, Bakshi and Madan (2006), Bollerslev et al. (2009), Carr and Wu (2009), and Jackwerth and Rubinstein (1996).

FIGURE 1

Option-Based Estimates of the Price of Coskewness Risk

In Figure 1, we plot the time series for the conditional physical and risk-neutral second moments (monthly in percentages) and the price of coskewness risk, multiplied by 100 for expositional convenience. The physical second moment is estimated using the heterogeneous autoregressive (HAR) model of Corsi (2009), and the risk-neutral second moment is constructed using the squared Volatility Index ($VIX^2$). The time-varying price of coskewness risk is equal to the spread between the physical and risk-neutral moments. The sample period is from Jan. 1986 to Dec. 2012.




The crucial difference between our approach and the one in KL (1976) and HS (2000) is that it explicitly utilizes the no-arbitrage restrictions on contingent-claims prices. This leads to a very simple estimator of the price of risk.[11]

## C. A First Look at the Option-Implied Price of Coskewness Risk

Figure 1 plots our option-based estimate of $\lambda_{\mathrm{OI},t}^{\mathrm{COSK}} = E_t^P\left(R_{m,t+1}^2\right) - E_t^Q\left(R_{m,t+1}^2\right)$, where $E_t^P\left(R_{m,t+1}^2\right)$ in Graph A is obtained using the heterogeneous autoregressive (HAR) model of Corsi (2009), and $E_t^Q\left(R_{m,t+1}^2\right)$ in Graph B is estimated by the squared value of the VIX. The details of the estimation will be provided in Section III. For now, note that the price of coskewness risk, $\lambda_{\mathrm{OI},t}^{\mathrm{COSK}}$, in Graph C exhibits interesting spikes surrounding the 1998 Long-Term Capital Management (LTCM) collapse, the WorldCom bankruptcy in 2002, the credit crisis in 2008, and bad economic news in Europe and the United States at the end of 2011. Both the risk-neutral and physical variance display spikes around those events, but the spikes in the physical variance are typically relatively smaller than the risk-neutral spikes, leading to the negative spikes in the price of coskewness risk. Note also that the new,

[11]Barras and Malkhozov (2016) find that the market variance risk premium measured in the equity market is similar to the variance risk premium estimated from options but reject the null hypothesis that they are identical. Our approach assumes that these risks have the same prices in different markets.

genuinely conditional estimates of the price of coskewness risk in Graph C have the theoretically expected negative sign in almost every month.

## III. Estimating the Price of Coskewness Risk

We now present estimates of the price of coskewness 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 coskewness 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 coskewness risk and discuss the differences between our new estimates and conventional regression-based estimates.

### A. 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 as our estimate for the risk-neutral variance. Using the VIX has a number of advantages. The VIX provides a very simple benchmark because the data are readily available from the Chicago Board Options Exchange (CBOE). 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 Jan. 1986–Dec. 2012. For existing articles that use the VIX squared as a proxy for the expected risk-neutral second moment with a 1-month horizon, see, for instance, Bollerslev et al. (2009). In the robustness analysis in Section IV.B, we use an alternative approach to compute the risk-neutral variance, following Bakshi and Madan (2000).

### B. Estimating the Physical Second Moment

Our benchmark implementation uses the so-called HAR model of Corsi (2009) estimated on daily, weekly, and monthly realized variances via

$$V_{t+1,t+K}^m = \phi_0 + \phi_1 V_{t-1,t}^m + \phi_2 V_{t-4,t}^m + \phi_3 V_{t-20,t}^m + \varepsilon_{t+1,t+K}^m, \tag{17}$$

where $V_{t+1,t+K}^m$ denotes the market index $K$-days-ahead integrated variance. In the previous equation, the variance terms are defined as

$$V_{s,s+\tau}^m = V_s^m + V_{s+1}^m + \cdots + V_{s+\tau}^m, \tag{18}$$

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

$$\begin{aligned} V_t^m = {} & \ln\left(S_t^{\text{HIGH}}/S_t^{\text{OPEN}}\right)\left[\ln\left(S_t^{\text{HIGH}}/S_t^{\text{OPEN}}\right) - \ln\left(S_t^{\text{CLOSE}}/S_t^{\text{OPEN}}\right)\right] \\ & + \ln\left(S_t^{\text{LOW}}/S_t^{\text{OPEN}}\right)\left[\ln\left(S_t^{\text{LOW}}/S_t^{\text{OPEN}}\right) - \ln\left(S_t^{\text{CLOSE}}/S_t^{\text{OPEN}}\right)\right], \end{aligned} \tag{19}$$

where $S_t^{\text{CLOSE}}$ ($S_t^{\text{OPEN}}$) is the close (open) price of the market index, measured by the Standard & Poor's (S&P) 500, and $S_t^{\text{HIGH}}$ ($S_t^{\text{LOW}}$) denotes the market index highest

TABLE 1

Option-Based Price of Coskewness Risk

In Table 1, we provide descriptive statistics for the physical and risk-neutral second moments and the price of coskewness risk. The physical second moment is estimated using the heterogeneous autoregressive (HAR) model of Corsi (2009), and the risk-neutral second moment is proxied by the squared Volatility Index (VIX$^2$). The time-varying price of coskewness 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 Jan. 1986 to Dec. 2012.

| | Physical Second Moment | Risk-Neutral Second Moment | Price of Coskewness Risk |
|---|---|---|---|
| Mean | 0.3034 | 0.4499 | −0.1464 |
| Std. dev. | 0.2168 | 0.4133 | 0.2289 |
| Skewness | 8.9863 | 3.4318 | −2.2645 |
| Kurtosis | 109.6172 | 19.1077 | 12.8347 |
| First-order autocorrelation | 0.4996 | 0.7525 | 0.4352 |

(lowest) price on day $t$.[12] The HAR model in equations (17)–(19) parsimoniously allows for a highly persistent dynamic in volatility and employs the intraday information available in our relatively long historical sample. We estimate the model using ordinary least squares (OLS) and a recursive 10-year (120-month) window. For related applications of high–low information in dynamic volatility models, see Azalideh, Brandt, and Diebold (2002), Chou (2005), and Brandt and Jones (2006).[13]

## C. Option-Based Estimates of the Price of Coskewness Risk

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

$$\widehat{\lambda}_{\mathrm{OI},t}^{\mathrm{COSK}} = \widehat{E}_t^P\left(R_{m,t+1}^2\right) - \widehat{E}_t^Q\left(R_{m,t+1}^2\right). \tag{20}$$

Table 1 reports descriptive statistics for the estimates of the moments and the price of risk. We discuss the robustness of our results to the choice of physical and risk-neutral variance models in Section IV.B. Figure 1 indicates that the coskewness price of risk is negative for almost all months. Table 1 shows that on average, it is equal to $-0.146$. This negative sign is consistent with theory, and several existing empirical articles document a negative price of coskewness risk as well (see, e.g., KL (1976), HS (2000)).

## D. Regression-Based Estimates

Table 2 reports results for three models based on cross-sectional regressions. The first model incorporates exposure to the market factor, and the second combines coskewness and market factors. The third model also includes the Fama–French (1993) size and book-to-market factors, as well as the momentum factor.

[12]We rescale $V_t^m$ to ensure its average equals the sample variance from daily close-to-close returns.

[13]Corsi (2009) and subsequent HAR articles 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.

We estimate loadings using the same window we use for the HAR model, 10 years of monthly returns, and subsequently run a cross-sectional regression for the next month.

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 data sets that are commonly used in the existing literature. We use portfolios formed on size and book-to-market, on size and momentum, on size and short-term reversal, and 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 (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). We also report on the data set that combines these four sets of test assets. This test uses 100 test assets.

Table 2 contains results for two sample periods: our benchmark sample (1986–2012) in Panel A and a somewhat shorter sample (1996–2012) in Panel B. We consider the shorter sample to enable comparisons with other measures of the option-based risk premium, which are not available for the 1986–2012 sample, in the robustness section later in the article.

Although the adjusted $R^2$s in Table 2 increase substantially when the coskew factor is included in the model, an important conclusion is that the estimate of the price of coskewness risk depends critically on the model and portfolios used in estimation. Focusing on the 1986–2012 sample in Panel A, the price of coskewness risk ranges from a significant $-0.157$ in the model with the Fama–French and momentum factors estimated on size and long-term-reversal portfolios to an insignificant 0.015 in the bivariate model estimated on size and short-term-reversal portfolios.[14]

There are some very important differences between our empirical results and regression-based estimates of the price of coskewness risk. The average of our newly proposed estimate of the price of coskewness risk in Table 1, which is equal to $-0.146$, is larger (in absolute value) than all but one of the estimates obtained using the regression approach in Section III.D. 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 IV.

The estimates in Figure 1 are encouraging because not only do we obtain unconditional estimates, but we also present genuinely *conditional* month-by-month estimates of the price of risk that have the theoretically expected sign in almost every month. Moreover, although there is no guarantee that these negative estimates for the price of coskewness 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 continue producing plausible negative estimates of the price of coskewness risk. Note that the regression-based approach cannot provide us with genuinely conditional month-by-month estimates. The focus of the Fama–MacBeth approach is to obtain unconditional estimates of the price of risk by averaging the results from the monthly cross-sectional regressions.

[14]For expositional convenience, we report the estimated prices of coskewness and cokurtosis risk, as well as the moments used to construct these estimates, in percentage terms (i.e., multiplied by 100).

TABLE 2

**Regression-Based Estimates of the Price of Coskewness Risk**

In Table 2, we show the results of cross-sectional Fama–MacBeth regressions using monthly returns. Each month, we estimate factor exposures using a 120-month rolling window of monthly returns from a time-series regression of excess portfolio returns on the factors. We then run a cross-sectional regression of returns on estimated exposures to obtain the price of risk, $\lambda$. We report the mean (in percentages) of the price-of-risk estimates and the Fama–MacBeth *t*-statistics with Newey–West correction for serial correlation, using 1 lag. We consider two periods, 1986–2012 and 1996–2012, and four sets of test assets, obtained using sorts on size and, respectively, book-to-market (BM), momentum (MOM), short-term reversal (STR), and long-term reversal (LTR).

| | 25 Size/BM | | | 25 Size/MOM | | | 25 Size/STR | | | 25 Size/LTR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| *Panel A. 1986–2012* | | | | | | | | | | | | |
| $\lambda^{0}$ | 1.338 | 1.408 | 0.849 | 0.182 | 0.008 | 0.592 | 0.574 | 0.471 | 0.369 | 0.389 | 0.490 | −0.208 |
| | 2.85 | 2.63 | 2.32 | 0.53 | 0.02 | 1.56 | 1.74 | 1.26 | 0.80 | 1.41 | 1.68 | −0.54 |
| $\lambda^{MKT}$ | −0.637 | −0.742 | −0.231 | 0.472 | 0.585 | 0.089 | 0.073 | 0.206 | 0.218 | 0.370 | 0.236 | 0.829 |
| | −1.27 | −1.37 | −0.56 | 1.11 | 1.33 | 0.22 | 0.17 | 0.44 | 0.46 | 1.12 | 0.68 | 1.92 |
| $\lambda^{HML}$ | | | 0.017 | | | 0.069 | | | −0.034 | | | 0.033 |
| | | | 0.10 | | | 0.35 | | | −0.18 | | | 0.15 |
| $\lambda^{SMB}$ | | | 0.333 | | | 0.200 | | | 0.497 | | | 0.542 |
| | | | 1.81 | | | 0.66 | | | 1.38 | | | 2.29 |
| $\lambda^{MOM}$ | | | 0.436 | | | 0.570 | | | −0.459 | | | −0.055 |
| | | | 0.89 | | | 2.01 | | | −0.96 | | | −0.13 |
| $\lambda^{COSK}$ | | −0.051 | −0.073 | | −0.108 | 0.010 | | 0.015 | 0.076 | | −0.080 | −0.157 |
| | | −0.65 | −1.13 | | −1.61 | 0.18 | | 0.19 | 1.19 | | −1.12 | −2.39 |
| Adj. $R^2$ | 16.01 | 31.14 | 44.99 | 13.78 | 27.26 | 56.22 | 14.37 | 32.73 | 51.11 | 7.84 | 21.86 | 43.75 |
| *Panel B. 1996–2012* | | | | | | | | | | | | |
| $\lambda^{0}$ | 1.121 | 1.163 | 0.967 | 0.595 | 0.478 | 0.727 | 0.251 | 0.322 | 0.381 | 0.700 | 0.748 | −0.055 |
| | 1.74 | 1.53 | 2.13 | 1.43 | 1.00 | 1.46 | 0.65 | 0.71 | 0.64 | 2.34 | 2.33 | −0.12 |
| $\lambda^{MKT}$ | −0.456 | −0.594 | −0.466 | 0.070 | 0.097 | −0.174 | 0.371 | 0.278 | 0.089 | 0.058 | −0.059 | 0.571 |
| | −0.66 | −0.79 | −0.89 | 0.13 | 0.16 | −0.32 | 0.70 | 0.48 | 0.15 | 0.13 | −0.14 | 1.08 |
| $\lambda^{HML}$ | | | 0.172 | | | 0.293 | | | 0.104 | | | 0.196 |
| | | | 0.69 | | | 1.10 | | | 0.39 | | | 0.63 |
| $\lambda^{SMB}$ | | | 0.366 | | | 0.103 | | | 0.298 | | | 0.658 |
| | | | 1.43 | | | 0.27 | | | 0.60 | | | 2.08 |
| $\lambda^{MOM}$ | | | 0.994 | | | 0.430 | | | −0.459 | | | 0.503 |
| | | | 1.37 | | | 1.02 | | | −0.70 | | | 0.98 |
| $\lambda^{COSK}$ | | −0.114 | −0.059 | | −0.113 | 0.014 | | −0.068 | 0.086 | | −0.105 | −0.124 |
| | | −1.13 | −0.83 | | −1.19 | 0.22 | | −0.73 | 1.21 | | −1.35 | −2.31 |
| Adj. $R^2$ | 15.84 | 29.94 | 46.08 | 16.33 | 28.95 | 59.47 | 15.39 | 32.60 | 53.70 | 10.39 | 24.10 | 47.18 |

The month-by-month estimates of the price of risk are noisy by design. We discuss these estimates in the Supplementary Material to highlight the differences.

In summary, a comparison of our newly proposed estimates of the price of coskewness 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 coskewness risk is $-0.146$ and indicates a role for coskewness that is large in magnitude compared with regression-based approaches.[15] Third, we consistently obtain negative estimates of the price of *conditional* coskewness risk in our approach. We therefore conclude that our approach appears to be economically appealing. We now investigate if these economically appealing results help with real-time investor decision making.

# IV. Comparing Predictive Performance

In this section, we compare the predictive performance of the option-implied and regression-based estimates of the coskewness price of risk.

## A. Comparing Cross-Sectional Predictive Performance

In this section, our goal is to compare the predictive performance of the option-implied and regression-based estimates of the price of coskewness risk from a cross-sectional perspective. We proceed as follows: For each portfolio $p$ in a given set of test assets, the 1-month-ahead expected return from coskewness risk is defined by

$$\text{COSK}_{t+1|t}^{p,k} = \lambda_{k,t}^{\text{COSK}} \beta_{p,t}^{\text{COSK}}, \tag{21}$$

where $\lambda_{k,t}^{\text{COSK}}$ is either the option-implied price of coskewness risk, which we denote $\lambda_{\text{OI},t}^{\text{COSK}}$, using the HAR and VIX$^2$ specification, or the regression-based price of risk, which we denote $\lambda_{\text{RB},t}^{\text{COSK}}$. In each case, the estimate is constructed using information through the current month $t$.

Portfolio $p$'s loading $\beta_{p,t}^{\text{COSK}}$ on the market return squared and $\beta_{p,t}^{\text{MKT}}$ are estimated via OLS regression including month $t$ returns. Based on the model predictions, we can define the monthly forecast error for each model by

$$\varepsilon_{t+1}^{p,k} = R_{p,t+1} - \lambda_{\text{RB},t}^{\text{MKT}} \beta_{p,t}^{\text{MKT}} - \text{COSK}_{t+1|t}^{p,k}, \tag{22}$$

where $R_{p,t+1}$ is the ex post realized portfolio and market excess returns next month. Note that because we are focused on comparing predictions from different estimates

[15]One potential concern is that when coskewness risk is motivated by a Taylor approximation, as in Section II.A, the quality of the Taylor approximation may affect the estimates of the price of risk (Garlappi and Skoulakis (2011), Skoulakis (2012)). We verified that for constant relative risk-aversion (CRRA) utility and the distribution of growth rates in our sample, a second-order Taylor approximation is unlikely to strongly bias the estimate of coskewness risk. In any case, such a bias would not explain the differences between the regression-based and option-implied approach.

of the coskewness price of risk, we subtract portfolio $p$'s expected return due to its market exposure. When evaluating the regression-based price of coskewness risk, $\lambda_{\mathrm{RB},t}^{\mathrm{MKT}}$ is estimated using a bivariate model that includes both market return and market return squared. When evaluating the option-implied price of coskewness risk, $\lambda_{\mathrm{RB},t}^{\mathrm{MKT}}$ is estimated independently of $\lambda_{\mathrm{RB},t}^{\mathrm{COSK}}$. This approach prevents measurement errors in $\lambda_{\mathrm{RB},t}^{\mathrm{COSK}}$ from spilling over into $\lambda_{\mathrm{RB},t}^{\mathrm{MKT}}$. To construct the forecast errors in equation (22), we need the loadings $\beta_{p,t}^{\mathrm{MKT}}$ and $\beta_{p,t}^{\mathrm{COSK}}$. We estimate the MKT and COSK loadings of each portfolio jointly by bivariate regressions where the factors are the market excess return and the market excess return squared. This ensures the consistency of our methodology with the SDF specified in Proposition 1.

When implementing equations (21) and (22), we have to decide on the rolling window to estimate the portfolio loadings $\beta_{p,t}^{\mathrm{COSK}}$ and $\beta_{p,t}^{\mathrm{MKT}}$, as well as the window used for the prices of risk. Given the genuinely conditional nature of the option-based price of risk, we expect it to do better with shorter windows compared with the regression-based approach. In our benchmark results, we report on a 12-month moving average for the price of risk and a 120-month window for the factor loadings. The Supplementary Material reports robustness results for alternative windows.

Armed with the monthly forecast errors of each model, we compute two measures of relative performance across risk-price-estimation methods. The first is a simple difference in mean squared error that exploits the information embedded in the cross section of portfolios in a given set of test assets. We refer to it as ΔMSE, defined by

$$\Delta\mathrm{MSE} = \frac{1}{T}\sum_{t=1}^{T}\Delta\mathrm{MSE}_t, \tag{23}$$

where the monthly difference in mean squared error is defined by

$$\Delta\mathrm{MSE}_t = \left(\frac{1}{25}\sum_{p=1}^{25}\left(\varepsilon_t^{p,\mathrm{RB}}\right)^2 - \left(\varepsilon_t^{p,\mathrm{OI}}\right)^2\right)\times 12\times 100. \tag{24}$$

Note that we multiply the cross-sectional average of the monthly difference in mean squared errors by 12 to annualize and by 100 to express it as a percentage. Note that a positive ΔMSE indicates a superior forecast performance of the option-implied estimate versus the regression-based estimate.

We also consider a relative $R^2$ measure, which we refer to as $\Delta R^2$, to evaluate the forecast performance across the risk-price estimates. We compute the average $\Delta R^2$ across portfolios (in percentages) using[16]

$$\Delta R^2 = \frac{1}{25}\sum_{p=1}^{25}\Delta R_p^2, \tag{25}$$

[16]We also report test results that combine the four sets of test assets. In this case, equations (24) and (25) must be adjusted by substituting 100 for 25.

where the $\Delta R^2$ of a given portfolio $p$ in a given set of test assets is given by

$$(26) \qquad \Delta R_p^2 = \left(1 - \frac{\sum_{t=1}^{T}\left(\varepsilon_t^{p,\mathrm{OI}}\right)^2}{\sum_{t=1}^{T}\left(\varepsilon_t^{p,\mathrm{RB}}\right)^2}\right) \times 100.$$

Similar to the previous $\Delta$MSE measure, the $\Delta R^2$ is signed and measures the forecast performance of the option model relative to the regression model. A large and positive $\Delta R^2$ again indicates the superior forecast performance of the option-implied forecast versus the regression-based forecast. In contrast, a negative $\Delta R^2$ indicates superior forecast performance of the regression-based price of risk relative to the option-implied price of risk.

Based on these two measures of performance, our goal is to conduct statistical inference about the superiority of one price-of-risk calculation relative to the other. Although statistical inference about $\Delta$MSE can easily be conducted using standard Fama–MacBeth methodology by exploiting the time series of the monthly $\Delta\mathrm{MSE}_t$, inference about the $\Delta R^2$ faces two important challenges. First, the two models we consider are nonnested, whereas asymptotic theories and critical values for out-of-sample $\Delta R^2$ are mostly available for the case of nested models. Second, our benchmark sample period of 1986–2012 is relatively short, which may cast doubts about the validity of standard asymptotic arguments. As a result, several of our empirical tests that follow on bootstrap methods for inference.

Table 3 contains the results for $\Delta$MSE and $\Delta R^2$. First consider Panel A. For each of the four sets of 25 test assets, the $\Delta$MSE is positive. The differences in mean squared error range from 0.325 to 0.081. They are significantly different from 0 at the 5% level in two of the four cases, according to the Newey–West (NW) $t$-statistic when using the Student's $t$-distribution. To benchmark the one-sided Student $t$ $p$-values, we also report bootstrapped $p$-values. For a given set of test assets, the $p$-value corresponds to the probability of observing a $\Delta$MSE under the null (when the forecast performance of both models is equal) that is at least as large as its sample value. To this end, we bootstrap 100,000 paths of option-implied and regression-based squared forecast errors under the null hypothesis to obtain the distribution of $\Delta$MSE. Based on that distribution, we then infer the proportion of paths having higher $\Delta$MSE than the value reported in Panel A. Table 3 shows that although the bootstrap and Student $t$ $p$-values are similar, the bootstrap indicates statistical significance in three of the four cases.[17] Panel A also contains a bootstrapped 90% confidence interval for the empirical distribution of $\Delta$MSE. Note that three of the four 5th-percentile bounds of the empirical distribution of $\Delta$MSE are larger than 0. In these cases, the forecasting ability of the option-implied price of risk dominates the regression-based price of risk more than 95% of the time.[18]

[17]Details on the bootstrap procedures can be found in Appendix A.3 of the Supplementary Material.

[18]Table 3 uses 12-month moving averages of the price of risk. In the Supplementary Material, we conduct an extensive robustness analysis with respect to this assumption. The option-implied price of risk consistently outperforms the regression-based price of risk, but more so for shorter windows. We also investigate robustness with respect to the window used to compute the risk exposures (120 months in Table 3).

TABLE 3

Coskewness Out-of-Sample Tests (1986–2012)

In Table 3, for each portfolio $p$ in a given set of test assets, we estimate model $k$'s monthly forecast error as

$$\varepsilon_{t+1}^{p,k} = R_{p,t+1} - \lambda_{\text{RB},t}^{\text{MKT}} \beta_{p,t}^{\text{MKT}} - \lambda_{k,t}^{\text{COSK}} \beta_{p,t}^{\text{COSK}},$$

where $\lambda_{\text{RB},t}^{\text{MKT}}$ is the 12-month moving average of the regression-based price of market risk, and $\beta_{p,t}^{F}$ is portfolio $p$'s loading to factor $F$ estimated by multivariate ordinary least squares (OLS) regression based on the most recent 120 months, including month $t$. $\varepsilon_{t+1}^{p,\text{OI}}$ is then constructed using $\lambda_{\text{OI},t}^{\text{COSK}}$, which is the 12-month moving average of the option-implied price of coskewness risk (estimated using the heterogeneous autoregressive (HAR) model of Corsi (2009) and the squared Volatility Index (VIX$^2$)), and $\varepsilon_{t+1}^{p,\text{RB}}$ is constructed using $\lambda_{\text{RB},t}^{\text{COSK}}$, which is the 12-month moving average of the regression-based price of risk. Each month, we calculate from these the cross-sectional difference in the mean squared error of the models' forecasts using all 25 portfolios $p$ in a given set of test assets according to

$$\Delta\text{MSE}_t = \left(\frac{1}{25}\sum_{p=1}^{25}\left(\varepsilon_t^{p,\text{RB}}\right)^2 - \left(\varepsilon_t^{p,\text{OI}}\right)^2\right) \times 12 \times 100,$$

where we multiply by 12 to annualize and by 100 to express the difference in percentages. We then compute the sample average of this measure for each test asset, $\Delta\text{MSE} = \sum_{t=1}^{T} \Delta\text{MSE}_t / T$. We also compute the percentage relative $R^2$ across portfolios according to

$$\Delta R^2 = \frac{1}{25}\sum_{p=1}^{25}\left(1 - \frac{\sum_{t=1}^{T}\left(\varepsilon_t^{p,\text{OI}}\right)^2}{\sum_{t=1}^{T}\left(\varepsilon_t^{p,\text{RB}}\right)^2}\right) \times 100.$$

Panel A reports the $\Delta$MSE for each set of test assets, the Newey–West (NW) $p$-value, the bootstrapped (BS) 1-sided $p$-value of $\Delta$MSE, and the 90% confidence bounds. Panel B reports the $\Delta R^2$ for each set of test assets, its BS $p$-value, and 90% confidence bounds. In the last column of each panel, we report the results obtained by using all 100 portfolios. Refer to Appendixes A.1 and A.2 in the Supplementary Material for further details on the estimation of $p$-values and confidence bounds. All BS results are based on 100,000 draws. The data are monthly, and the sample period is from Jan. 1986 to Dec. 2012.

| | 25 Size/BM | 25 Size/MOM | 25 Size/STR | 25 Size/LTR | All |
|---|---|---|---|---|---|
| *Panel A. Difference in Mean Squared Error* | | | | | |
| $\Delta$MSE | 0.325 | 0.134 | 0.233 | 0.081 | 0.193 |
| NW $p$-value | 0.30% | 5.35% | 0.30% | 15.61% | 0.40% |
| BS $p$-value | 0.15% | 3.37% | 0.14% | 13.58% | 0.16% |
| BS 5th-percentile bound | 0.171 | 0.015 | 0.112 | −0.040 | 0.091 |
| BS 95th-percentile bound | 0.492 | 0.255 | 0.359 | 0.202 | 0.298 |
| *Panel B. Relative $R^2$* | | | | | |
| $\Delta R^2$ | 5.80 | 2.61 | 4.69 | 1.71 | 3.70 |
| BS $p$-value | 0.03% | 2.62% | 0.05% | 17.33% | 0.13% |
| BS 5th-percentile bound | 3.18 | 0.32 | 2.32 | −1.29 | 1.65 |
| BS 95th-percentile bound | 8.45 | 4.82 | 6.98 | 4.77 | 5.74 |

Panel B of Table 3 reports various statistics on the $\Delta R^2$ computed across the 25 portfolios for each set of test assets. The $\Delta R^2$s are economically large across test assets and range from 1.71% for size/long-term reversal (LTR) to 5.80% for size/book-to-market (BM). Most importantly, the one-sided $p$-values of the $\Delta R^2$ are very small in three out of four cases, showing that the option-implied price of coskewness risk significantly outperforms the regression-based estimate when using $\Delta R^2$.

The last column in Panels A and B of Table 3 reports results for all 100 test assets simultaneously. Both the $\Delta$MSE and $\Delta R^2$ results are economically large and statistically significant.

We conclude from Table 3 that the cross-sectional predictive evidence in favor of our new option-based estimates of the price of coskewness risk is strong, economically large, and statistically significant. Note that the option-implied estimates of the price of risk do not depend on the estimated loadings, but the predictions in Table 3 use these loadings. Estimation error in the loadings will therefore affect predictive performance. When comparing the regression-based and option-based estimates, we implicitly assume that these errors equally affect both predictions. We leave a more detailed study of the implications of errors in coskewness loadings for future research.

## B. Robustness

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

In our benchmark results, we simply use the square of the VIX as our estimate for the risk-neutral variance. In the robustness analysis, we use the Bakshi et al. (BKM) (2003) replication approach to obtain an estimate of the risk-neutral second moment. The BKM approach is based on a continuum of out-of-the-money call and put options, which is approximated using cubic spline interpolation techniques. We implement the BKM approach using data on S&P 500 index options from Option-Metrics for the period Jan. 1996–Dec. 2012. See the Supplementary Material for more details on the implementation.

We investigate two alternative approaches to the HAR approach for modeling the conditional physical variance. The first is a simple autoregressive (AR) model on realized variances. Like the HAR model, it is estimated using OLS and a 10-year recursive window. The one-step-ahead forecast of the physical second moment is estimated from the following monthly regression:

(27) $$V_t^m = a_0 + a_1 V_{t-1}^m + u_{V,t},$$

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.

We also consider the Heston (1993) stochastic volatility model in which the market index return follows

(28) $$\frac{dS_t}{S_t} = \mu dt + \sqrt{V_t^m} dW_{S,t},$$

and the instantaneous variance dynamic is

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

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 the Euler discretized version of this model using maximum likelihood and the particle filter on a 10-year moving window of daily returns. To ensure consistency with our measure of the risk-neutral variance, we generate the forecasts of the physical 1-month total variance

(30) $$\widehat{V}_{t+1,t+21}^m = \theta + (V_t - \theta) \frac{1 - \exp(-(21/252)\kappa)}{\kappa(21/252)},$$

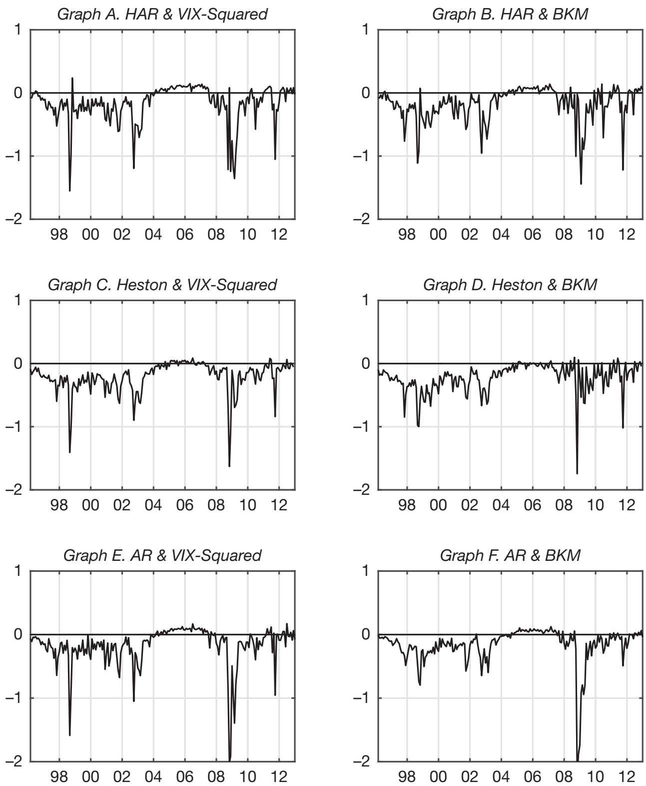
from the model at the end of every month.

Figure 2 plots the six different option-implied prices of coskewness risk constructed by combining one of the two risk-neutral variances (VIX or BKM) with one of the three physical variances (HAR, Heston, and AR). Graph A contains

FIGURE 2

Various Option-Implied Estimates of the Price of Coskewness Risk Using Alternative Models for Physical and Risk Neutral Variance (1996–2012)

In Figure 2, we plot the price of coskewness risk, defined as the physical variance minus the risk-neutral variance. The physical variance is either from the heterogeneous autoregressive (HAR) model of Corsi (2009) estimated on daily high–low returns, from the Heston (1993) model estimated on returns, or from an autoregressive (AR) model estimated on realized monthly variances from daily squared returns. The risk-neutral variance is either the square of the Volatility Index (VIX) from the Chicago Board Options Exchange (CBOE) or based on Bakshi et al. (BKM) (2003).




our benchmark HAR/VIX estimate from Figure 1 for reference. The six graphs in Figure 2 demonstrate that the estimated prices of coskewness risk are quite robust across methods and are highly correlated.

In the Supplementary Material, we present cross-sectional prediction results using the different coskew estimates in Figure 2. The estimates of the average price of risk and the out-of-sample tests in Table 3 are very robust. The option-implied coskew risk estimates deliver a superior fit to the cross section of equity portfolio ex post coskewness returns, regardless of which risk-neutral and physical variance estimate is used. The Supplementary Material also reports on the 1986–2007 sample period to provide insights into the impact of the turbulent 2008–2009 financial crisis period on the performance of the option-implied price of risk. The results are robust and, if anything, even stronger in favor of the option-implied estimates than the benchmark results in Table 3.

We conclude that the performance of the new estimate of the price of coskewness risk is not driven by the 2008–2009 financial crisis nor by the particular implementation of the physical and risk-neutral moments.

# V. Measuring Higher-Moment Risk

This section generalizes and extends the results from Section II. First, we consider the estimation of the price of cokurtosis risk using index option prices. Subsequently, we discuss how our approach can be generalized beyond coskewness and cokurtosis to more general risks that are nonlinear in the market return.

## A. Measuring Cokurtosis Risk

A natural extension of the quadratic pricing kernel in equation (5) is the cubic pricing kernel studied by Dittmar (2002), given by

$$(31)\qquad m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P\left(R_{m,t+1}^2\right)\right) + b_{3,t}\left(R_{m,t+1}^3 - E_t^P\left(R_{m,t+1}^3\right)\right).$$

A cubic pricing kernel is consistent with investors' preferences for higher-order moments, specifically skewness and kurtosis (i.e., $U'''(\cdot) \neq 0$ and $U''''(\cdot) \neq 0$). 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 cokurtosis risk, in addition to covariance risk and coskewness risk. As explained by Dittmar (2002), kurtosis measures the likelihood of extreme values, and cokurtosis 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 cokurtosis risk, meaning that the price of cokurtosis risk should be positive. See Guidolin and Timmermann (2008) and Scott and Horvath (1980) for a more detailed discussion. Similar to coskewness risk, cokurtosis risk has been defined in various ways in previous articles. For instance, Ang et al. (2006) measure cokurtosis risk using

$$\left(E^P\left[(R_j - \overline{R}_j)(R_m - \overline{R}_m)^3\right]\right) \Big/ \left(\sqrt{E^P\left[(R_j - \overline{R}_j)^2\right]}\left(E^P\left[(R_m - \overline{R}_m)^2\right]\right)^{3/2}\right),$$

and Guidolin and Timmermann (2008) use $\text{cov}\left(R_j, R_m^3\right)$. In this article, we measure cokurtosis risk by the return's loading on the cubic market return $R_m^3$. We denote the cokurtosis loading of a stock by $\beta_{j,t}^{\text{COKU}}$.

The following proposition presents the estimator for the cokurtosis price of risk and the cross-sectional pricing restrictions:

*Proposition 2.* In the absence of arbitrage opportunities, if the SDF has the following form:

$$(32)\qquad m_{t+1} = a_t + b_{1,t}\left(R_{m,t+1} - E_t^P(R_{m,t+1})\right) + b_{2,t}\left(R_{m,t+1}^2 - E_t^P\left(R_{m,t+1}^2\right)\right) + b_{3,t}\left(R_{m,t+1}^3 - E_t^P\left(R_{m,t+1}^3\right)\right),$$

then the cross-sectional restriction on stock returns is

$$E_t^P\left(R_{j,t+1}\right) - R_{f,t} = \lambda_t^{\mathrm{MKT}} \beta_{j,t}^{\mathrm{MKT}} + \lambda_t^{\mathrm{COSK}} \beta_{j,t}^{\mathrm{COSK}} + \lambda_t^{\mathrm{COKU}} \beta_{j,t}^{\mathrm{COKU}}, \tag{33}$$

where $\beta_{j,t}^{\mathrm{MKT}}$, $\beta_{j,t}^{\mathrm{COSK}}$, and $\beta_{j,t}^{\mathrm{COKU}}$ 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, $\lambda_t^{\mathrm{MKT}}$, and coskewness risk, $\lambda_t^{\mathrm{COSK}}$, are

$$\lambda_t^{\mathrm{MKT}} = E_t^P(R_{m,t+1}) - R_{f,t}, \tag{34}$$

$$\lambda_t^{\mathrm{COSK}} = E_t^P\left(R_{m,t+1}^2\right) - E_t^Q\left(R_{m,t+1}^2\right), \tag{35}$$

and the price of cokurtosis risk, $\lambda_t^{\mathrm{COKU}}$, is

$$\lambda_t^{\mathrm{COKU}} = E_t^P\left(R_{m,t+1}^3\right) - E_t^Q\left(R_{m,t+1}^3\right), \tag{36}$$

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

*Proof.* The structure of the proof largely follows the proof of Proposition 1. Given equation (32), applying equation (14) for $\Psi_{t+1} \equiv R_{m,t+1}$ as in Proposition 1, we recover equation (34), and applying equation (14) for $\Psi_{t+1} \equiv R_{m,t+1}^2$, we recover equation (35). In addition, applying equation (14) for $\Psi_{t+1} \equiv R_{m,t+1}^3$, we obtain equation (36). This again uses the results that a return that is also a factor has a loading of 1 onto itself and 0 on the other factors. □

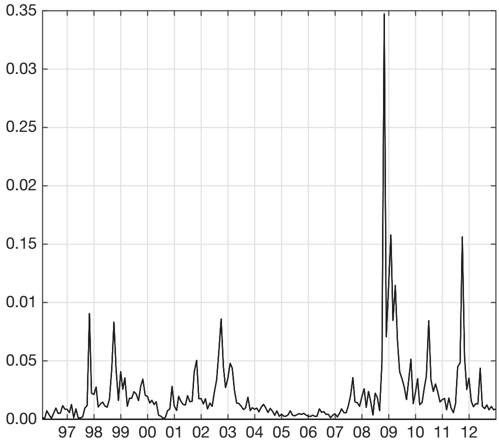
Proposition 2 shows that the price of cokurtosis risk is equal to the spread between the market physical and risk-neutral third moments. Existing empirical evidence (see, e.g., Bakshi et al. (2003)) indicates that the risk-neutral distribution for the market return is more left-skewed than the physical distribution, therefore suggesting a positive price of cokurtosis risk. This is consistent with theory, as explained by Dittmar (2002).

In our empirical implementation, we estimate the risk-neutral third moment using the method of Bakshi and Madan (2000). We implement the Bakshi and Madan (2000) approach using data on S&P 500 index options from OptionMetrics for the period Jan. 1996–Dec. 2012. The Supplementary Material contains the details. It is well known that estimation of the physical third moment is more 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. The unconditional third-moment estimate for monthly S&P 500 returns during the 1986–2012 period is not statistically different from 0 at conventional confidence levels, and moreover, it is very small compared with the estimates of risk-neutral moments in our sample. We therefore present and discuss different estimates of the price of cokurtosis risk. In our benchmark implementation, we set the physical third moment equal to 0. Setting the physical third

FIGURE 3

**Option-Based Estimates of the Price of Cokurtosis Risk**

In Figure 3, we plot the price of cokurtosis risk, multiplied by 100 for expositional convenience. We report on the benchmark case where the physical third moment is set equal to 0. The risk-neutral moment is estimated using the model-free approach of Bakshi and Madan (2000) and Bakshi et al. (2003). The sample period is from Jan. 1996 to Dec. 2012.




moment to 0 is preferable to using noisy estimates. We also report on the price of cokurtosis risk using two alternative estimates for the physical third moment: a constant third moment computed using daily data and a fully dynamic physical third moment, estimated using a version of the dynamic moment model of Jondeau and Rockinger (2003) and described in the Supplementary Material. The resulting estimates of the price of cokurtosis risk are 0.022, 0.0214, and 0.0116, respectively.[19]

Figure 3 depicts the time series of the price of cokurtosis risk for our benchmark implementation, which is simply the negative of the risk-neutral third moment. Consistent with theory, the price of cokurtosis risk in Figure 3 is positive throughout the period. Existing empirical articles have also documented positive prices of cokurtosis risk. See, for instance, Ang et al. (2006), who find that stocks with higher cokurtosis earn higher returns. The Supplementary Material discusses additional empirical results. The model with cokurtosis risk performs well in predicting the cross section of portfolio returns when using the price of risk estimated from option-based risk premia, but its performance somewhat depends on the measure of the physical third moment.

## B. More General SDFs

We now present the general case that nests the results in Propositions 1 and 2. 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. However, although the empirical focus of this article is on cokurtosis

[19]Recall that for expositional convenience, these estimates of the moments and the prices of coskewness and cokurtosis risk are reported as percentages.

and especially coskewness risk, it is important to note that 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 also includes more complex nonlinear relationships, such as, for instance, measures of downside risk as in Ang et al. (2006). The following proposition presents the result for the general case:

*Proposition 3*. In the absence of arbitrage opportunities, if the SDF has the following form:

(37)
$$m_{t+1} = a_t + \sum_i b_{i,t}\left(G_i(R_{m,t+1}) - E_t^P(G_i(R_{m,t+1}))\right) + \sum_l c_{l,t}\left(f_{l,t+1} - E_t^P(f_{l,t+1})\right),$$

where $G_i(R_{m,t+1})$ is a nonlinear function of the market return, and $f_{l,t+1}$ is the realization of risk factor $l$, then the cross-sectional pricing restriction for stock returns is

(38)
$$E_t^P(R_{j,t+1}) - R_{f,t} = \sum_i \lambda_t^i \beta_{j,t}^i + \sum_l \gamma_t^l \beta_{j,t}^l,$$

where $\beta_{j,t}^i$ and $\beta_{j,t}^l$ are from the projection of asset returns on $G_i(R_{m,t+1})$ 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 $G_i$, of the market return, $\lambda_t^i$, is

(39)
$$\lambda_t^i = E_t^P(G_i(R_{m,t+1})) - E_t^Q(G_i(R_{m,t+1})),$$

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

*Proof.* The structure of the proof is again similar to the proof of Proposition 1. Given equation (37), then applying equation (14) for $\Psi_{t+1} \equiv G_i(R_{m,t+1})$, we obtain equation (39). □

Proposition 3 shows that the reward for exposure to any nonlinear function $G(\cdot)$ 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. Fortunately, the spanning approach of Bakshi and Madan (2000) allows us to use options to replicate the risk-neutral expectation of any twice-differentiable function.

## VI. Conclusion

We propose an alternative strategy for estimating the price of possibly nonlinear exposures to market risk that avoids some shortcomings inherent in the cross-sectional and Fama–MacBeth 2-step regression approaches. The key difference between our approach and existing methods is that we explicitly impose the pricing restrictions on both stocks and contingent claims. We study two important applications of our general approach: The price of coskewness risk in our framework corresponds to the spread between the physical and risk-neutral second moments. The price of cokurtosis risk is similarly given by the spread between the physical and risk-neutral third moments.

We find that the price of coskewness risk has the theoretically expected negative sign in almost every month, and the price of cokurtosis risk has the theoretically expected positive sign in almost every month. Our approach thus provides genuinely conditional estimates of the price of risk at monthly or even higher frequencies. An out-of-sample analysis of factor models with coskewness and cokurtosis risk indicates that the new estimates of the price of risk improve the models' performance compared with regression-based estimates.

Several extensions of our work are possible. First, Bollerslev et al. (2009) demonstrate that variance risk is priced when the representative agent has Epstein–Zin–Weil preferences, and Bali and Zhou (2016) use this setup to investigate the cross section of stock returns. In their framework, the market variance risk premium is a pricing factor, whereas in our article, it is a price of risk. A further investigation of how these results are related would be of interest. Second, Chabi-Yo (2012) and Maheu, McCurdy, and Zhao (2013) study the relative importance of (co)skewness and volatility risk, and our estimates can be used to revisit this question. Third, for computing the price of cokurtosis risk, it may prove interesting to use the CBOE SKEW index, which, like the VIX, is readily available. Improved modeling of the physical third moment may also lead to improvements in the estimates of the price of cokurtosis risk. Fourth, although the focus of our new approach is on improving the measurement of the price of risk, we worry that the estimated loadings we use in the analysis may be noisy and affect the performance of the newly proposed prices of risk. Improved estimation of loadings 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 coskewness and cokurtosis risk that takes into account the uncertainty in the various steps involved in the computation.

## Supplementary Material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S002210902000023X.

## References

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

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

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

Bakshi, G.; N. Kapadia; and D. Madan. "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options." *Review of Financial Studies*, 16 (2003), 101–143.
Bakshi, G., and D. Madan. "Spanning and Derivative Security Valuation." *Journal of Financial Economics*, 55 (2000), 205–238.
Bakshi, G., and D. Madan. "A Theory of Volatility Spreads." *Management Science*, 52 (2006), 1945–1956.
Bakshi, G.; D. Madan; and G. Panayotov. "Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels." *Journal of Financial Economics*, 97 (2010), 130–154.
Bali, T. G., and R. F. Engle. "The Intertemporal Capital Asset Pricing Model with Dynamic Conditional Correlations." *Journal of Monetary Economics*, 57 (2010), 377–390.
Bali, T. G., and H. Zhou. "Risk, Uncertainty, and Expected Returns." *Journal of Financial and Quantitative Analysis*, 51 (2016), 707–735.
Bansal, R., and S. Viswanathan. "No-Arbitrage and Arbitrage Pricing: A New Approach." *Journal of Finance*, 48 (1993), 1231–1262.
Barone-Adesi, G. "Arbitrage Equilibrium with Skewed Asset Returns." *Journal of Financial and Quantitative Analysis*, 20 (1985), 299–313.
Barras, L., and A. Malkhozov. "Does Variance Risk Have Two Prices? Evidence from the Equity and Option Markets." *Journal of Financial Economics*, 121 (2016), 79–92.
Barro, R. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics*, 121 (2006), 823–866.
Bollerslev, T.; J. Marrone; L. Xu; and H. Zhou. "Predicting Stock Returns with Variance Risk Premia: Statistical Inference and International Evidence." *Journal of Financial and Quantitative Analysis*, 49 (2014), 633–661.
Bollerslev, T.; G. Tauchen; and H. Zhou. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22 (2009), 4463–4492.
Bollerslev, T., and V. Todorov. "Tails, Fears, and Risk Premia." *Journal of Finance*, 66 (2011), 2165–2211.
Brandt, M., and C. Jones. "Volatility Forecasting with Range-Based EGARCH Models." *Journal of Business and Economic Statistics*, 24 (2006), 470–486.
Campbell, J. Y. "Intertemporal Asset Pricing Without Consumption Data." *American Economic Review*, 83 (1993), 487–515.
Campbell, J. Y. "Understanding Risk and Return." *Journal of Political Economy*, 104 (1996), 298–345.
Carr, P., and L. Wu. "Variance Risk Premia." *Review of Financial Studies*, 22 (2009), 1311–1341.
Chabi-Yo, F. "Conditioning Information and Variance Bounds on Pricing Kernels with Higher-Order Moments: Theory and Evidence." *Review of Financial Studies*, 21 (2008), 181–231.
Chabi-Yo, F. "Pricing Kernels with Stochastic Skewness and Volatility Risk." *Management Science*, 58 (2012), 624–640.
Chabi-Yo, F. "Real-Time Distribution of Stochastic Discount Factors." Working Paper, University of Massachusetts at Amherst (2018).
Chabi-Yo, F.; D. Leisen; and E. Renault. "Aggregation of Preferences for Skewed Asset Returns." *Journal of Economic Theory*, 154 (2014), 453–489.
Chang, B.-Y.; P. Christoffersen; and K. Jacobs. "Market Skewness Risk and the Cross-Section of Stock Returns." *Journal of Financial Economics*, 107 (2013), 46–68.
Chaudhuri, R., and M. Schroder. "Monotonicity of the Stochastic Discount Factor and Expected Option Returns." *Review of Financial Studies*, 28 (2015), 1462–1505.
Chen, H.; S. Joslin; and N.-K. Tran. "Rare Disasters and Risk Sharing with Heterogeneous Beliefs." *Review of Financial Studies*, 25 (2012), 2189–2224.
Chou, R. "Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range." *Journal of Money, Credit and Banking*, 37 (2005), 561–582.
Cochrane, J. *Asset Pricing*. Princeton, NJ: Princeton University Press (2005).
Conrad, J.; R. Dittmar; and E. Ghysels. "Ex Ante Skewness and Expected Stock Returns." *Journal of Finance*, 68 (2013), 85–124.
Constantinides, G., and D. Duffie. "Asset Pricing with Heterogeneous Consumers." *Journal of Political Economy*, 104 (1996), 219–240.
Corsi, F. "A Simple Approximate Long Memory Model of Realized Volatility." *Journal of Financial Econometrics*, 7 (2009), 174–196.
Cremers, M.; M. Halling; and D. Weinbaum. "Aggregate Jump and Volatility Risk in the Cross-Section of Stock Returns." *Journal of Finance*, 70 (2015), 577–614.
Dittmar, R. F. "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns." *Journal of Finance*, 57 (2002), 369–403.
Drechsler, I., and A. Yaron. "What's Vol Got to Do with It." *Review of Financial Studies*, 24 (2011), 1–45.

Driessen, J.; P. Maenhout; and G. Vilkov. “The Price of Correlation Risk: Evidence from Equity Options.” *Journal of Finance*, 64 (2009), 1377–1406.
Fama, E. F., and K. R. French. “Common Risk Factors in the Returns on Stocks and Bonds.” *Journal of Financial Economics*, 33 (1993), 3–56.
Fama, E. F., and J. D. MacBeth. “Risk, Return and Equilibrium: Empirical Tests.” *Journal of Political Economy*, 81 (1973), 607–636.
Gabaix, X. “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance.” *Quarterly Journal of Economics*, 127 (2012), 645–700.
Garlappi, L., and G. Skoulakis. “Taylor Series Approximations to Expected Utility and Optimal Portfolio Choice.” *Mathematics and Financial Economics*, 5 (2011), 121–156.
Ghysels, E. “On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?” *Journal of Finance*, 53 (1998), 549–573.
Ghysels, E.; A. Plazzi; and R. Valkanov. “Why Invest in Emerging Markets? The Role of Conditional Return Asymmetry.” *Journal of Finance*, 71 (2016), 2145–2192.
Golec, J., and M. Tamarkin. “Bettors Love Skewness, Not Risk, at the Horse Track.” *Journal of Political Economy*, 106 (1998), 205–225.
Gospodinov, N.; R. Kan; and C. Robotti. “Misspecification-Robust Inference in Linear Asset Pricing Models with Irrelevant Risk Factors.” *Review of Financial Studies*, 27 (2014), 2139–2170.
Gourio, F. “Disaster Risk and Business Cycles.” *American Economic Review*, 102 (2012), 2734–2766.
Guidolin, M., and A. Timmermann. “International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences.” *Review of Financial Studies*, 21 (2008), 889–935.
Han, B., and Y. Zhou. “Variance Risk Premiums and Cross-Section of Stock Returns.” Working Paper, University of Texas at Austin (2011).
Harvey, C. R., and A. Siddique. “Conditional Skewness in Asset Pricing.” *Journal of Finance*, 55 (2000), 1263–1295.
Heston, S. “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options.” *Review of Financial Studies*, 6 (1993), 327–343.
Jackwerth, J., and M. Rubinstein. “Recovering Probability Distributions from Option Prices.” *Journal of Finance*, 51 (1996), 1611–1631.
Jagannathan, R., and Z. Wang. “An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression.” *Journal of Finance*, 53 (1998), 1285–1309.
Jondeau, E., and M. Rockinger. “Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements.” *Journal of Economic Dynamics and Control*, 27 (2003), 1699–1737.
Julliard, C., and A. Ghosh. “Can Rare Events Explain the Equity Premium Puzzle?” *Review of Financial Studies*, 25 (2012), 3037–3076.
Kan, R., and C. Zhang. “Two-Pass Tests of Asset Pricing Models with Useless Factors.” *Journal of Finance*, 54 (1999), 204–235.
Kelly, B., and H. Jiang. “Tail Risk and Asset Prices.” *Review of Financial Studies*, 27 (2014), 2841–2871.
Kleibergen, F. “Tests of Risk Premia in Linear Factor Models.” *Journal of Econometrics*, 149 (2009), 149–173.
Kraus, A., and R. H. Litzenberger. “Skewness Preference and the Valuation of Risk Assets.” *Journal of Finance*, 31 (1976), 1085–1100.
Leland, H. “Beyond Mean-Variance: Performance Measurement of Portfolios Using Options or Dynamic Strategies.” Working Paper, University of California, Berkeley (1997).
Lim, K.-G. “A New Test of the Three-Moment Capital Asset Pricing Model.” *Journal of Financial and Quantitative Analysis*, 24 (1989), 205–216.
Liu, J.; J. Pan; and T. Wang. “An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks.” *Review of Financial Studies*, 18 (2005), 131–164.
Maheu, J.; T. McCurdy; and X. Zhao. “Do Jumps Contribute to the Dynamics of the Equity Premium?” *Journal of Financial Economics*, 110 (2013), 457–477.
Martin, I. “What Is the Expected Return on the Market?” *Quarterly Journal of Economics*, 132 (2017), 367–433.
Mitton, T., and K. Vorkink. “Equilibrium Underdiversification and the Preference for Skewness.” *Review of Financial Studies*, 20 (2007), 1255–1288.
Patton, A., and A. Timmermann. “Monotonicity in Asset Returns: New Tests with Applications to the Term Structure, the CAPM, and Portfolio Sorts.” *Journal of Financial Economics*, 98 (2010), 605–625.
Rietz, T. “The Equity Risk Premium: A Solution.” *Journal of Monetary Economics*, 22 (1988), 117–131.
Rogers, L., and S. Satchell. “Estimating Variance from High, Low and Closing Prices.” *Annals of Applied Probability*, 1 (1991), 504–512.

Rubinstein, M. "The Fundamental Theory of Parameter-Preference Security Valuation." *Journal of Financial and Quantitative Analysis*, 8 (1973), 61–69.
Schneider, P. "Generalized Risk Premia." *Journal of Financial Economics*, 116 (2015), 487–504.
Schneider, P.; C. Wagner; and J. Zechner. "Low-Risk Anomalies?" *Journal of Finance* 75 (2020), 2673–2718.
Scott, R. C., and P. A. Horvath. "On the Direction of Preference for Moments of Higher Order than the Variance." *Journal of Finance*, 35 (1980), 915–919.
Shanken, J. "On the Estimation of Beta Pricing Models." *Review of Financial Studies*, 5 (1992), 1–33.
Skoulakis, G. "On the Quality of Taylor Approximations to Expected Utility." *Applied Financial Economics*, 22 (2012), 863–876.
Wachter, J. "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 68 (2013), 987–1035.