Do Stock Markets Price Expected Stock Skewness? New Evidence from Quantile Regression based Skewness Forecasts*

Kevin Aretz†
Manchester Business School
kevin.aretz@mbs.ac.uk

Eser Arisoy‡
Université Paris-Dauphine
eser.arisoy@dauphine.fr

This Draft: December 7, 2015

Abstract

We propose an estimator of an asset's future return skewness based on density forecasts estimated from quantile regressions. The estimator is unbiased and efficient, and it can easily be adapted to forecast skewness over any conceivable return interval. Using Neuberger's (2012) *realized* skewness, we show that skewness forecasts derived from the estimator outperform skewness forecasts used in other studies. Notwithstanding, they do not condition stock returns, neither independently nor when (optimally) combined with the other forecasts. All in all, our evidence casts doubt on whether stock markets price expected stock skewness, as suggested by a growing literature.

Key words: Asset pricing; ex-ante skewness; realized skewness; quantile regression models.

JEL classification: G11, G12, G15.

†Corresponding author, Accounting and Finance Division, Booth Street West, Manchester M15 6PB, UK, tel.: +44(0)161 275 6368, e-mail: <kevin.aretz@mbs.ac.uk>.

‡Department of Finance, Place du Maréchal de Lattre de Tassigny, Cedex 16, Paris 75775, France, tel: +33(0) 144 054 360, e-mail: <eser.arisoy@dauphine.fr>.

*We are indebted to Sofiane Aboura, David Allan, Sonia Konstantinidi, Alex Kostakis, Anthony Neuberger, Peter Pope, and participants at the 2014-Amundi Asset Management Workshop in Paris for suggestions.

Do Stock Markets Price Expected Stock Skewness?

New Evidence from Quantile Regression based Skewness Forecasts*

This Draft: December 7, 2015

Abstract

We propose an estimator of an asset's future return skewness based on density forecasts estimated from quantile regressions. The estimator is unbiased and efficient, and it can easily be adapted to forecast skewness over any conceivable return interval. Using Neuberger's (2012) *realized* skewness, we show that skewness forecasts derived from the estimator outperform skewness forecasts used in other studies. Notwithstanding, they do not condition stock returns, neither independently nor when (optimally) combined with the other forecasts. All in all, our evidence casts doubt on whether stock markets price expected stock skewness, as suggested by a growing literature.

Key words: Asset pricing; ex-ante skewness; realized skewness; quantile regression models.

JEL classification: G11, G12, G15.

1 Introduction

Recent studies suggest that assets whose future returns are expected to be more positively skewed have lower expected returns than assets whose future returns are expected to be less positively skewed. The theoretical literature derives a negative relation between an asset's expected return and its expected skewness from models featuring endogenous investor beliefs (Brunnermeier et al. (2007)), heterogeneity in investors' skewness preferences (Mitton and Vorkink (2007)), or prospect-theory preferences (Barberis and Huang (2008)). The empirical literature finds a negative relation between proxies meant to capture the expected skewness of stock returns over investors' holding horizons ("expected stock skewness") and the cross-section of stock returns (Boyer et al. (2010); Bali et al. (2011); and Conrad et al. (2014)).

We offer further tests of how expected stock skewness relates to the cross-section of stock returns. We do so because prior empirical work relies on somewhat adhoc proxies to capture expected stock skewness. For example, Boyer et al. (2010) use an ordinary least-squares (OLS) forecast of the realized skewness of daily returns. Conrad et al. (2014) use the fitted value from a logit model predicting whether a stock has an above 100% future-return. Whether such indirect proxies are well suited to capture expected stock skewness over the long horizons that investors care about is neither theoretically obvious nor empirically shown.[1]

Our contribution to the literature is threefold. Our first contribution is to introduce a methodology that allows us to generate expected stock skewness estimates over any conceivable return interval. To implement the methodology, we run quantile regressions of the h month-ahead return on a set of lagged predictor variables. Next, we calculate stock-specific estimates of the first three return moments from the fitted regression values and then use them to

[1] Take Boyer et al.'s (2010) expected stock skewness proxy as example. Unless daily returns are identically distributed, their proxy captures the *average* skewness of daily returns over the period over which realized skewness is calculated. However, even if daily returns were identically distributed, skewness would still not scale nicely with time unless daily returns were also independent. Thus, unless we also assume independence, their proxy would not be informative about the skewness of long-ahead returns. This insight is problematic because, ultimately, it is the skewness of *long-ahead* returns that *long-term* investors care about. It is precisely for this reason that Neuberger (2012, p.3424) asks "why asset prices in an economy with well-capitalized long-term investors should be heavily influenced by the characteristics of short-term returns."

calculate stock-specific skewness coefficients. In addition to being adaptable to different return intervals, our methodology does not suffer from low precision due most stocks having short time-series of non-overlapping low-frequency returns. Neither does it suffer from survivorship bias because we are not forced to exclude stocks that are delisted over the quantile-regression estimation period. Finally, the methodology relies on precisely estimated quantiles — and not on imprecisely estimated means (see Koenker (2005) for the details).

We forecast the skewness of one month-, one year-, and five year-ahead stock returns. The forecasts are derived from a limited set of fitted quantiles: the first, fifth, tenth, 25th, 50th, 75th, 90th, 95th, and 99th. The fitted quantiles are modeled as linear functions of a comprehensive set of pre-determined firm fundamentals that prior literature suggests to be related to stock skewness, as, for example, market size, historical volatility, and share turnover. Using rolling window estimations, we ensure that the skewness forecasts could have been calculated in calendar time. Preliminary tests suggest that the fitted quantiles of all but the longest-ahead return are well calibrated. More specifically, we show that the proportions of future returns that fall below the different estimated quantiles align with expectations. For example, around 10% of the one-year ahead future returns fall below the tenth-quantile estimate.

We also calculate a more naive, but well-known quantile-regression based skewness forecast. This forecast is the distance between the third quartile and the median minus the distance between the first quartile and the median, scaled by the distance between the two outer quantiles (Kim and White (2004) and Konstaninidi and Pope (2015)).

Our second contribution is to compare the forecasting power of the quantile-regression based skewness forecasts with those of the skewness forecasts used in the prior literature. While, until only recently, it would have been hard to do so, Neuberger (2012) derives a realized skewness estimator which is similar to the well-known realized volatility estimator studied in, for example, Andersen and Bollerslev (1998). The similarity arises because the realized skewness estimator also aggregates up high-frequency (daily) data to estimate realized skewness over lower-frequency intervals. However, in contrast to the realized volatility estimator, the realized skewness estimator requires both stock data and option data as inputs. We seem to be first in applying the realized skewness estimator to single stocks.

Using realized skewness as forecast target, we run unbiasedness tests, portfolio formation exercises, and optimal forecast-combination regressions to compare forecasting power. The quantile-regression based skewness forecast derived from the fitted moments comes out best in all tests. The naive quantile-regression forecast does not accurately capture the skewness of short-ahead returns, but does better in capturing the skewness of longer-ahead returns. Of the skewness forecasts used in prior studies, Conrad et al.'s (2014) logit model prediction is a strong contender to the moment-implied quantile-regression based skewness forecast. The other skewness forecasts are often disappointing in capturing realized skewness.

As final contribution, we return to the question of whether expected stock skewness is priced. We conduct portfolio formation exercises and Fama-MacBeth (1973) regressions using the various skewness forecasts as pricing variables. We also use optimal or equally-weighted combinations of the skewness forecasts, using only information available at the time to derive the optimal combinations. Results show that the quantile-regression based forecasts are never significantly priced. For example, the quintile portfolio of stocks with the highest values for the moment-implied quantile-regression based twelve-month ahead forecast has an only 5.4% lower mean return per annum than the quintile portfolio of stocks with the lowest values (t-stat: –0.84). The skewness combinations also never attract significant premia.

Confirming prior work, some of the non-quantile-regression based skewness forecasts are significantly priced. Boyer et al.'s (2010) realized skewness forecast is significantly negatively priced, but only in tests in which observations are value-weighted. Bali et al.'s (2011) maximum return is also significantly negatively priced, but only in tests in which observations are equally-weighted. Due to our sample period omitting data from the 1960s and 1970s, Conrad et al.'s (2014) logit model-based skewness forecast never attracts a significant premium.

Our work contributes to studies examining whether expected skewness is priced. Scott and Horvath (1980) show that expected-utility investors with preferences that do not depend on wealth levels have a positive preference for odd moments (such as skewness) and a negative preference for even moments. Because adding more assets to a portfolio can decrease expected skewness, Simkowitz and Beedles (1978) and Conine and Tamarkin (1981) show that positive skewness preferences can lead investors to hold under-diversified portfolios.

Nevertheless, expected returns do not necessarily depend on expected skewness. Rubinstein (1973) shows that investors that care about skewness but no other higher moments choose portfolios in such a way that expected asset returns are proportional to their covariances and co-skewness with wealth. Assuming monetary separation,[2] expected asset returns are also proportional to their return covariances and co-skewness with aggregate wealth. Under these assumptions, expected skewness does not play a role over and above co-skewness.

To establish expected skewness as separate pricing factor, monetary separation must not hold. Mitton and Vorkink (2007) prevent monetary separation from holding by allowing for heterogeneity in investors' skewness preferences. Brunnermeier et al. (2007) and Barberis and Huang (2008) prevent monetary separation from holding by using non-expected utility preferences. However, not even under the assumptions in the above papers is expected skewness always priced. For example, Barberis and Huang (2008) show that, unless the skewed asset in their model is expected to be extremely skewed, a CAPM-type of equilibrium prevails. Thus, whether investors price expected skewness is ultimately an empirical question.

Boyer et al. (2010), Bali et al. (2011), and Conrad et al. (2014) show that their proxies for expected stock skewness are negatively related to the cross-section of stock returns. For example, Boyer et al. (2010) report that the quintile portfolio holding the stocks with the lowest values for their skewness forecast has a mean return of 1.189% per month, while the quintile portfolio holding the stocks with the highest values has a mean return of 0.515% per month. The spread is highly statistically significant. Using a skewness forecast derived from

[2]Monetary separation implies that investors' optimal portfolios are a combination of the risk-free asset and a portfolio of the risky assets. Cass and Stiglitz (1970) show that a sufficient condition for monetary separation is that the ratio of the second derivative of investors' utility functions with respect to wealth to the corresponding first derivative is linear in wealth, with a homogeneous slope coefficient.

the strike price and the assumption that the stock price is log-normal, Boyer and Vorkink (2014) also find a negative skewness premium in the stock-options market.

Conrad et al. (2013) document that *risk-neutral* stock-skewness forecasts derived from the methodology of Bakshi et al. (2003) are negatively related to stock returns. However, neither Xing et al. (2010), who use the option smirk, nor Stilger et al. (2014), who follow Conrad et al. (2013) in using Bakshi et al.'s (2003) methodology, are able to confirm these results. Instead, they both report a positive relationship between the two variables. Even if expected risk-neutral skewness were priced, it would not necessarily follow that expected (physical) skewness is priced. For example, Rubinstein's (1976) model suggests that an increase in the covariance between an asset's payoff and consumption (systematic risk) lowers risk-neutral skewness, raises the expected return, but does *not* change the physical skewness. Thus, the pricing of expected risk-neutral skewness could be a pure systematic-risk effect.

Our evidence should caution us to not prematurely accept the idea that stock markets price expected skewness. Using quantile-regression based skewness forecasts that outperform the forecasts used in studies empirically supporting a negative expected stock skewness-stock return relationship, we fail to find any evidence suggesting such a relationship.

Our article is structured as follows. Section 2 describes how we use quantile regressions to calculate skewness forecasts. It also offers details on Neuberger's (2012) realized-skewness methodology, gives an overview of the skewness forecasts used in prior studies, and describes our data sources. Section 3 offers the forecast comparisons, the optimal combination tests, and the asset pricing tests. Section 4 summarizes and concludes.

2 Methodology & Data

2.1 The Quantile-Regression Based Skewness Forecasts

2.1.1 Methodology

Quantile regressions model the conditional quantiles of an endogenous variable using a linear function of exogenous variables (Koenker and Bassett (1978); Koenker (2005); and Angrist and Pischke (2009)). To see how this works, denote by $F(y)$ the cumulative density function of the random variable y. Also, denote by $q_\tau(y) = \inf\{y : F(y) \geq \tau\}$ random variable y's τth quantile, with $0 < \tau < 1$. The quantile regression model can then be written as:

$$q_\tau(y_i)|\mathbf{X}_i = \mathbf{X}_i'\beta_\tau, \tag{1}$$

where y_i is the ith observation of the endogenous variable, $\mathbf{X}_i$ is a vector of exogenous variables, and β_τ is a parameter vector. Equation (1) assumes that the exogenous variables exert a locally monotonic effect on the quartile. To wit, raising the value of an exogenous variable from, say, minus one to zero has the same effect as raising its value from zero to one. Despite this restriction, quantile regression models generate flexible quantile estimates, owing to the fact that the parameter vector β_τ is allowed to vary over the modeled quantiles.

Defining the regression residual ϵ_i as $y_i - \mathbf{X}_i'\beta_\tau$, an estimate of the parameter vector β_τ can be obtained by minimizing the following "tick"-loss function $L_\tau(\epsilon_i)$:

$$\begin{aligned} L_\tau(\epsilon_i) &= (\tau \mathbf{1}\{\epsilon_i \geq 0\} + (1-\tau)\mathbf{1}\{\epsilon_i < 0\})|\epsilon_i| && (2) \\ &= (\tau \mathbf{1}\{\epsilon_i \geq 0\} - (1-\tau)\mathbf{1}\{\epsilon_i < 0\})\epsilon_i && (3) \\ &= (\tau - \mathbf{1}\{\epsilon_i < 0\})\epsilon_i. && (4) \end{aligned}$$

where $\mathbf{1}\{\epsilon_i \geq 0\}$ is an indicator function equal to one if $\epsilon_i \geq 0$ and else zero, and $\mathbf{1}\{\epsilon_i < 0\}$ is defined accordingly. Conceptually speaking, minimizing $L_\tau(\epsilon_i)$ estimates the τth quantile by assigning different penalties to overpredicting ($\epsilon_i < 0$) and underpredicting ($\epsilon_i \geq 0$) the endogenous variable. For example, if τ is 0.25, overestimating the endogenous variable is three times more costly than underestimating it. To compensate, the model optimally chooses the probability of an overprediction to be three times smaller than the probability of an underprediction. In other words, optimality requires that $3 \times \text{Prob}(y > q_\tau(y_i)|\mathbf{X}_i) = \text{Prob}(y \leq q_\tau(y_i)|\mathbf{X}_i)$. Thus, the optimal $q_\tau(y_i)$ is the conditional first quartile of y_i.

In theory, we could estimate β_τ by setting the directional derivatives of $\sum_i L_\tau(\epsilon_i)$ with respect to β_τ to non-negative values. In practice, the problem is usually reformulated as:

$$\min_{\beta_\tau, \mathbf{u_n}, \mathbf{v_n}} \{\tau \mathbf{1}'\mathbf{u_n} + (1-\tau)\mathbf{1}'\mathbf{v_n} | \epsilon = \mathbf{u_n} - \mathbf{v_n}\}, \tag{5}$$

where ϵ is vector of residuals, and $\mathbf{u_n}$ and $\mathbf{v_n}$ are vectors of slack variables. The slack variables need to obey: (i) $u_i \geq 0$, (ii) $v_i \geq 0$, and (iii) $u_i \times v_i = 0$. The new problem can be solved using standard linear programming techniques (see Koenker (2005) for details).

The variance-covariance matrix of the β_τ estimate, $\mathbf{V}_n$, is calculated using:

$$\mathbf{V}_n = \hat{s}_\tau^2 \tau (1-\tau) \left(\sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i' \right), \tag{6}$$

where $\hat{s}_\tau$ is equal to the inverse of an estimate of the value of the endogenous variable's density function evaluated at the τth quantile. The pseudo R-squared is given by one minus the sum of the weighted regression-residuals derived from the fitted quantiles scaled by the sum of the weighted regression-residuals derived from a quantile-regression model using only a constant as exogenous variable. The weight applied to the residuals is equal to τ for positively valued residuals, and it is equal to $(1-\tau)$ for negatively valued residuals.

As a next step, we transform the conditional fitted quantiles into moment estimates. To do so, we order the fitted quantiles from lowest to highest and then collect them in the vector $\hat{Q}(y_i)$. We denote the jth element of $\hat{Q}(y_i)$ by $\hat{Q}_j(y_i)$, where $j = 1, \ldots, J$ and J is the total number of elements in the vector. Assuming that the endogenous variable y is uniformly distributed between two consecutive quantiles, it holds that:

$$\hat{E}[y_i|\hat{Q}_{j-1}(y_i) \leq y_i < \hat{Q}_j(y_i)] = \frac{\hat{Q}_{j-1}(y_i) + \hat{Q}_j(y_i)}{2}, \tag{7}$$

$$\hat{E}[y_i^2|\hat{Q}_{j-1}(y_i) \leq y_i < \hat{Q}_j(y_i)] = \frac{\hat{Q}_{j-1}(y_i)^2 + \hat{Q}_{j-1}(y_i) \times \hat{Q}_j(y_i) + \hat{Q}_j(y_i)^2}{3}, \tag{8}$$

$$\hat{E}[y_i^3|\hat{Q}_{j-1}(y_i) \leq y_i < \hat{Q}_j(y_i)] = \frac{\hat{Q}_{j-1}(y_i)^3 + \hat{Q}_{j-1}(y_i)^2 \times \hat{Q}_j(y_i) + \hat{Q}_{j-1}(y_i) \times \hat{Q}_j(y_i)^2 + \hat{Q}_j(y_i)^3}{4}, \tag{9}$$

where $\hat{E}$ is an estimate of the expectation operator. Using an approximation to the law of total probability, the unconditional expectations can be estimated from:

$$\hat{E}[y_i^n] = \sum_{j=2}^{J} \frac{F^{-1}(\hat{Q}_j(y_i)) - F^{-1}(\hat{Q}_{j-1}(y_i))}{F^{-1}(\hat{Q}_J(y_i)) - F^{-1}(\hat{Q}_1(y_i))} \hat{E}[y_i^n|\hat{Q}_{j-1}(y_i) \leq y_i < \hat{Q}_j(y_i)], \tag{10}$$

where n is one, two, or three. Equation (10) approximates the integral over the conditional nth moment defining the unconditional nth moment. However, because there is positive probability mass below the lowest and above the highest fitted quantile, the approximation is rescaled by $F^{-1}(\hat{Q}_J(y_i)) - F^{-1}(\hat{Q}_1(y_i))$. Using the fitted moments obtained from Equation (10), we calculate the moment-implied quantile-regression based skewness coefficient as:

$$QRSkew_i = E\left[\left(\frac{y_i - E[y_i]}{\sigma_{y,i}}\right)^3\right] = \frac{E[y_i^3] - 3E[y_i]\sigma_{y,i}^2 - E[y_i]^3}{(E[y_i^2] - E[y_i]^2)^{\frac{3}{2}}}, \tag{11}$$

where $\sigma_{y,i}$ is the standard deviation of y_i, defined as $(E[y_i^2] - E[y_i]^2)^{\frac{1}{2}}$.

Kim and White (2004) and Konstantinidi and Pope (2015) use another approach to convert the fitted quantiles obtained from quantile regressions into skewness coefficients. Using their approach, the skewness coefficient is calculated as follows:

$$NaiveQRSkew_i = \frac{[(\hat{q}_{75}(y_i) - \hat{q}_{50}(y_i)) - (\hat{q}_{50}(y_i) - \hat{q}_{25}(y_i))]}{\hat{q}_{75}(y_i) - \hat{q}_{25}(y_i)}, \tag{12}$$

where, for convenience, the equation omits the dependence of the quantiles on $\mathbf{X}_i$.

Prior work suggests that quantile regressions are well suited to fit stock return quantiles. For example, Cenesizoglu and Timmermann (2008) show that quantile regressions produce efficient forecasts of the conditional return density of the S&P 500, in-sample and out-of-sample. They also show that the quantile-regression based forecasts contain useful information over and above other variables. In particular, they significantly improve market timing- and option investment-strategies relative to other S&P 500 stock-return density forecasts.

2.1.2 Implementation Details

We calculate the quantile-regression based skewness forecasts as follows. The endogenous variable in the quantile regressions is the h-month ahead return. For example, when $h = 12$, the endogenous variable is the return compounded over months $t+1$ to $t+12$. An exception occurs when a stock is delisted between months t and $t+h$. If this happens, the endogenous variables is the return compounded up until (and including) the delisting month. Doing so, the estimations do not exclude delisted stocks and thus avoid survivorship bias. In the empirical tests, we set h equal to one, twelve (one year), and 60 (five years).

Following other studies, we forecast the skewness of total and idiosyncratic returns. To calculate idiosyncratic returns, we run stock-specific time-series regressions of the stock's return on the market return minus the risk-free rate, SMB, HML, and MOM:

$$R_{i,t} = \alpha_i + \beta_{i,t}^{MKT}(R_{mkt,t} - r_{f,t}) + \beta_{i,t}^{SMB} R_{SMB,t} + \beta_{i,t}^{HML} R_{HML,t} + \beta_{i,t}^{MOM} R_{MOM,t} + \epsilon_{i,t}, \quad (13)$$

where $R_{i,t}$ is the return of stock i in month t, $R_{mkt,t} - r_{f,t}$ is the market return minus the risk-free rate, $R_{SMB,t}$ the return of a size spread portfolio, $R_{HML,t}$ the return of a book-to-market spread portfolio, and $R_{MOM,t}$ is the return of a one-year past return spread portfolio. See Kenneth French's website for more details. α_i, $\beta_{i,t}^{MKT}$, $\beta_{i,t}^{SMB}$, $\beta_{i,t}^{HML}$, and $\beta_{i,t}^{MOM}$ are free parameters, and $\epsilon_{i,t}$ is the residual. We run the time-series regression over the prior 20 years of monthly data for each stock with more than 60 months of data. We compound $\alpha_i + \epsilon_{i,t}$

over months t to $t+h$ and use it as endogenous variable in the regressions.

In line with Boyer et al. (2010), we include historical volatility, historical skewness, market capitalization ("size"), the past return ("momentum"), share turnover, a NASDAQ dummy, and industry dummies among the exogenous variables in the regressions. Harvey and Siddique (2000), Chen et al. (2001), and Conrad et al. (2014) show that a higher *Hist. Volatility* predicts a more positive future skewness, possibly due to it indicating the existence important growth options or due to the limited-liability feature of stocks. Chen et al. (2001) document that a worse past performance, as signalled by a lower *Momentum* value, predicts a more positive future skewness. Hong and Stein (2003) report that a higher *TradingVolume* forecasts a more negative future skewness. Boyer et al. (2010) add *Size*, *NASDAQ*, and the industry dummies, presumably to control for differences between differently sized firms that are operating in different industries and that are traded on different types of stock exchanges.

We add to the exogenous variables suggested by Boyer et al. (2010) the book-to-market ratio, share issuances, accruals, asset growth, and profitability. Chen et al. (2001) report that a lower *BookToMarket* value forecasts a more negative future skewness. The theoretical work of Barberis and Huang (2008) suggests that higher *ShareIssuances* and *AssetGrowth* values signal the existence of important growth options generating positive future skewness. Because accounting conservatism implies that bad news are recognized on a timelier basis than good news, firms following more conservative accounting practices tend to have lower *Accruals* (earnings minus cash flows) and thus possibly more positively skewed future returns. Finally, Campbell et al. (2008) and Conrad et al. (2014) show that a lower *Profitability* helps in identifying financially distressed firms with more right skewed future returns.

We follow the above studies in constructing the exogenous variables. More details about the construction of the exogenous variables is offered in Table A.1 in the Appendix.[3]

[3]Similar to Boyer et al. (2010), we do not winsorize the market variables used in the quantile regression models. However, following Fama and French (2008), we winsorize the variables wholly or partially reliant on accounting data, such as *BookToMarket*, *ShareIssuance*, *Accruals*, *AssetGrowth*, and *Profitability*. We also winsorize the control variables used in the asset pricing tests. We always winsorize at the 0.5th and 99.5th percentiles, using percentiles calculated separately by month.

We estimate the quantile regressions using rolling windows of data. At the end of each month t in our sample period, we extract the data over months $t-239$ to t from our full data sample. For each stock and each month t^* within the rolling window (except the most recent h months), we create the h month-ahead return by compounding a stock's returns over months t^*+1 to t^*+h. We deal with stocks that are delisted between months t^*+1 to t^*+h as described above. We next estimate panel quantile regressions of the h month-ahead returns on the exogenous variables measured in month t^*, fitting the first, fifth, tenth, 25th, 50th, 75th, 90th, 95th, and 99th quantiles of the h month ahead-return density.[4] We combine the regression estimates with the values of the exogenous variables at the end of the rolling window (i.e., at the end of month t) to calculate the fitted quantiles. Using the fitted quantile, we then calculate the quintile-regression based skewness forecasts. Thus, the skewness forecasts are conditional on only information available to investors in month t. Also, they are constructed to capture the skewness of a stock's return over months $t+1$ to $t+h$.

Our main tests consider the one-month ($QRSkew_{t+1}$), one-year ($QRSkew_{t+1,t+12}$), and five-year ahead ($QRSkew_{t+1,t+60}$) moment-implied quantile-regression based skewness forecasts calculated from raw returns. They also investigate the naive counterparts of these forecasts ($NaiveQRSkew_{t+1}$, $NaiveQRSkew_{t+1,t+12}$, and $NaiveQRSkew_{t+1,t+60}$, respectively). In unreported tests, we further study the above six variables calculated from idiosyncratic returns or Boyer et al.'s (2010) exogenous variables. Results are qualitatively unaffected.

2.1.3 Other Skewness Forecasts

Skewness forecasts used in the prior literature are often *indirect* forecasts that cannot easily be tailored to returns calculated over different intervals. Boyer et al.'s (2010) skewness forecast ($OLSSkew$) is the fitted value from a cross-sectional OLS regression of the realized skewness of daily returns on firm fundamentals. The daily returns span the period from months $t+1$ to

[4] Adding more extreme quantiles (e.g., the 0.1th and 99.9th quantiles) does not add to forecasting power, probably because the more extreme quantiles are estimated with very low precision.

$t + 60$. The firm fundamentals include the variables referred to above and are observed at the end of month t. Boyer et al. (2010) combine the regression results with the firm fundamental values in month $t + 60$ to forecast skewness over the period starting with month $t + 61$.

Conrad et al. (2014) analyze the fitted value from a logit regression of a dummy variable equal to one if a stock's return over months $t + 1$ to $t + 12$ exceeds 100% and else zero, on firm fundamentals measured in month t (*LogitSkew*). In line with us, they include historical volatility, historical skewness, market size, momentum, and share turnover among the firm fundamentals.[5] In addition, they use sales growth, company age, and asset tangibility. Sales growth is the log of the ratio of current sales to one-year lagged sales; age is the number of years since a stock first appeared in CRSP; and asset tangibility is the ratio of gross property, plant, and equipment to total assets.[6] The logit model is estimated over recursive windows starting with June 1951 and ending with June of year $t - 12$. The model only considers June observations. The logit estimates are combined with the values of the firm fundamentals from June of year t to May of year $t + 1$ to capture stock skewness over the next twelve month.

Bali et al. (2011) use a stock's maximum daily return during month $t - 1$ to capture its propensity to produce a lottery-like (and thus right skewed) future return (*MaxSkew*).

Another possible skewness forecast is a stock's historical skewness, calculated using either daily or even higher frequency data over some past period (*HistoricalSkew*). However, because Harvey and Siddique (1999) find that skewness is not persistent over time, only few studies use historical skewness to forecast future skewness.[7] We use historical skewness calculated from daily data over the prior month to complement the other skewness forecasts.

We follow Boyer et al. (2010), Bali et al. (2011), and Conrad et al. (2014) in calculating

[5] However, different from us, they calculate skewness and volatility using returns centered around zero and over months $t - 2$ to t, and they use log returns in their skewness calculations. Also, they calculate share turnover as the average past six-month turnover minus the average past 18-month turnover.

[6] Conrad et al. (2014) winsorize all continuous analysis variables at the fifth and 95th percentiles. Somewhat surprisingly, Boyer et al. (2010) do not indicate that they winsorize their analysis variables.

[7] An exception is Amaya et al. (2015), who calculate historical stock skewness from intra-day data. They show that their variable is negatively related to next week's stock return, but do not offer evidence that the negative relationship arises because intra-day historical skewness forecasts future skewness.

their skewness forecasts. Consistent with the quantile-regression based skewness forecasts, their forecasts also only rely on data available to investors in calendar time.

2.2 Neuberger's (2012) Realized Skewness

2.2.1 Methodology

Assuming that stock prices are martingales, Neuberger (2012) shows that an estimate of the realized skewness of a stock's *log* return can be obtained from:

$$RealizedSkew_{t,t+T} = \frac{\sum_{t=1}^{T}\left(3\Delta\upsilon^{E}(t)(e^{\Delta s(t)}-1)+K(\Delta s(t))\right)}{\left(\sum_{t=1}^{T}2(e^{\Delta s(t)}-1-\Delta s(t))\right)^{3/2}}, \tag{14}$$

where $K(\Delta s(t)) = 6(\Delta s(t)e^{\Delta s(t)} - 2e^{\Delta s(t)} + \Delta s(t) + 2)$, $\Delta s(t)$ is the change in the log stock price from period $t-1$ to t, and $\Delta\upsilon(t)$ is the change in the stock's "entropy variance" from period $t-1$ to t. The stock's entropy variance in period t, υ_t^E, is defined as:

$$\upsilon_t^E = \mathbb{E}_t[2(s_T - s_t)e^{(s_T - s_t)} - e^{(s_T - s_t)} + 1], \tag{15}$$

where $\mathbb{E}_t$ is the expectation operator conditional on period t information, and s_t is the log stock price in period t. Using the insight that any financial claim can be replicated using a portfolio of plain-vanilla options, we can write υ_t^E as:

$$\upsilon_t^E = \int_{K=0}^{F}\frac{2P(K)}{FK}dK + \int_{K=F}^{\infty}\frac{2C(K)}{FK}dK, \tag{16}$$

where $C(K)$ and $P(K)$ are, respectively, the values of plain-vanilla European call and put options with strike price K and an invariant maturity date, and F is the price of a forward contract with the same maturity date as the options (Carr and Wu (2007, 2009)).

When stock prices are not martingales, the realized skewness in Equation (14) is upward biased (downward biased) if the stock price is positively (negatively) correlated with the variance risk premium. The expected bias can be written as:

$$6\mathbb{E}\left[\sum_{u=1}^{T-1}\sum_{t=u}^{T-1}\left(\frac{\Delta E(u)-\Delta_0(t)\Delta S(u)}{S_0}\frac{\Delta S(t+1)}{S_t}+\frac{\Delta S(u)}{S_0}\frac{\Delta E_{t+1}-\Delta(t)\Delta S(t+1)}{S_t}\right)\right], \quad (17)$$

where $\Delta E(u)$ is the change in the value of an entropy contract from period $u-1$ to u, where the "entropy contract" pays $S_T \ln S_T$ at maturity, $\Delta S(u)$ is the change in the stock price from period $u-1$ to u, S_t is the stock price in period t, and $\Delta(t) = 1 + \ln S_t + \frac{1}{2}\upsilon_t^E$.

Neuberger's (2012) realized skewness captures the skewness of *log* returns. In contrast, the quantile-regression skewness based forecasts capture the skewness of *discrete* returns. Thus, even if the quantile-regression based skewness forecasts have a high forecasting power, we still expect their mean levels to exceed those of the realized skewness estimates. Notwithstanding, since the skewness of discrete returns is strongly positively correlated with the skewness of log returns in the cross-section,[8] we still expect the quantile-regression based skewness forecasts to accurately predict a stock's ranking according to realized skewness if the forecasting power of the quantile regression-based skewness forecasts is high.

2.2.2 Implementation Details

We use daily stock and option data to calculate *RealizedSkew*. Because there are no European options on single stocks, we are forced to use American options to estimate a stock's entropy variance. While this choice is inconsistent with Equation (16), we mitigate resulting inaccuracies by selecting American options with a relatively short time-to-maturity (and thus a low early-exercise premium). In particular, we use options whose time-to-maturity is closest to, but exceeds two weeks. We eliminate stock option-day observations with fewer than 20 strike prices, and stock options with less than one year of data. We use a cubic regression model of implied volatility on strike prices and time-to-maturities to smooth the implied

[8]For example, we find a mean cross-sectional correlation of 0.97 between the skewness of discrete daily returns and the skewness of log daily returns in CRSP data between 1963 to 2010.

volatility surface. The model coefficients are estimated on a weekly basis. We winsorize the implied volatility estimates at the first and 99th percentiles, calculated by date.

We approximate the integral in Equation (16) using:

$$\begin{aligned} \upsilon_t^E \approx & \sum_{i=1}^{n} \frac{2P(K_i)}{FK_i}(K_i - K_{i-1}) + \frac{2P(K_{n+1})}{FK_{n+1}}(F - K_n) \\ & + \frac{2C(K_{n+1})}{FK_{n+1}}(K_{n+1} - F) + \sum_{i=n+2}^{m} \frac{2C(K_i)}{FK_i}(K_i - K_{i-1}), \end{aligned} \tag{18}$$

where $0 = K_0, K_1, K_2, \ldots, K_n, F, K_{n+1}, \ldots, K_m$ is a set of prices ranked in ascending order, with K_i being a strike price and F the forward price. We calculate the forward price as the stock price multiplied by the exponential of the risk-free rate of return times the time-to-maturity. We use the smoothed implied volatilities derived above to calculate the sum from the lowest to the highest available strike price, using the step sizes found in the data. We use the implied volatility of the option with the lowest (highest) available strike price to approximate the sum up to (starting from) this strike price, choosing the distance between the two lowest (highest) available strike prices as step size. Plugging the daily stock return and the daily change in υ_t^E into Equation (14), we obtain realized skewness estimates. We calculate realized skewness over the next one month, one year, and five years.

2.3 Data Sources

Market variables are from CRSP and accounting variables from COMPUSTAT. Option data are from OptionMetrics. Data on the market return, SMB, HML, and MOM are from Kenneth French's website.[9] *OLSSkew* is from Brian Boyer's website.[10] Our tests exclude stocks from the financial (SIC codes: 6000-6999) and utilities industry (4900-4949). We also exclude stocks

[9]The URL address is: <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>. We thank Eugene Fama and Kenneth French for making their data available.

[10]The URL address is: <http://marriottschool.net/emp/boyer/>. Although the sample period studied in their paper is December 1987 to December 2007, the data available from Brian Boyer's website extends to December 2010. We thank Brian Boyer and his co-authors for making their data available.

with a negative book value or a price below $1 in month $t-1$. We use data over the 1968-2010 period to estimate the quantile regression models. However, following Boyer et al. (2010), we run the asset pricing tests over the December 1987 to December 2010 period.

3 Empirical Results

3.1 Quantile Regression Estimates

3.1.1 Coefficient Estimates and Fitted Quantiles

Table 1 shows the results of quantile regressions fitting the first, fifth, tenth, 25th, 50th, 75th, 90th, 95th, and 99th quantiles of stocks' one-month ahead returns. The quantile regressions are estimated over all 20-year windows whose final month lies in the December 1987-December 2010 period. For each fitted quantile and exogenous regression variable, the table reports the mean coefficient (in bold) and the fraction of coefficients that are statistically significant at the 90% confidence level or better over the estimation windows (in parenthesis). For each fitted quantile, it also reports the mean R-squared. Quantile regressions of returns measured over different intervals (e.g., over the next one or five years) produce results similar to those in Table 1. For the sake of brevity, we thus do not report their results.

Table 1 suggests that most firm fundamentals help in explaining the density of stocks' one-month ahead returns. Particularly powerful are *Hist. Volatility*, *Size*, *Momentum*, *Profitability*, and *Turnover*, which attract significant coefficients in all nine quantile regression models over (almost) all estimation windows. *BookToMarket*, *ShareIssuances*, and *AssetGrowth* are more important for fitting the left tail of the return density, while *Accruals* is more important for fitting its body (i.e., returns around the median). Neither *Hist. Skew* nor *NASDAQ* play pivotal roles. Mean R-squareds range from zero to close to 20%. In line with Cenesizoglu and Timmermann (2008), mean R-squareds are larger for the models fitting the outer-quantiles, suggesting it is easier to fit the tails than the body of the return density.

Looking at how each exogenous variable's coefficients vary over the nine quantile regression models reveals interesting patterns, hinting at the variable's association with skewness. As an example, the coefficients of *Hist. Volatility* increase monotonically over the fitted quantiles, rising from 3.6 for the first quantile to 10.1 for the 99th quantile. As a result, an increase in *Hist. Volatility* increases the higher quantiles more than the lower quantiles, generating more right skewness. In comparison, decreases in *Size*, *BookToMarket*, or *Profitability* or increases in *Turnover* produce outward shifts both in the left tail and in the right tail. However, in case of *Size* and *Profitability*, the effect on the right tail dominates the effect on the left tail, generating more right skewness. In case of *BookToMarket* and *Turnover*, the effect on the left tail dominates the effect on the right tail, generating more left skewness.

To further examine how the firm fundamentals are related to expected stock skewness, Figure 1 plots the fitted one-month ahead return densities of the average stock in several extreme firm-fundamental deciles (one and ten). The average stock's firm-fundamental values are calculated by averaging the firm fundamentals first by decile and month and then by decile alone. We combine the average stock's firm-fundamental values with the mean coefficients in Table 1 to construct the stock's return density. Small and unprofitable stocks have extremely right skewed returns. Value stocks with a lot of share issuances or a low asset growth also have right skewed returns, but to a somewhat lesser degree. Given small and value stocks' extremely right skewed returns in Figure 1, their tendency to produce high future returns would be surprising if expected stock skewness were negatively priced.

3.1.2 Calibration of the Fitted Quantiles

As a next step, we investigate how well the fitted quantiles of the one-month, one-year, and five-year ahead return are calibrated. For each fitted quantile, we thus calculate the fraction of future returns that lie below the fitted quantile. If the fitted quantiles are well calibrated, we expect the fraction to be close to τ. For example, when looking at the fifth quantile of the one-year ahead return, we expect about 5% of the returns over months $t+1$ to $t+12$

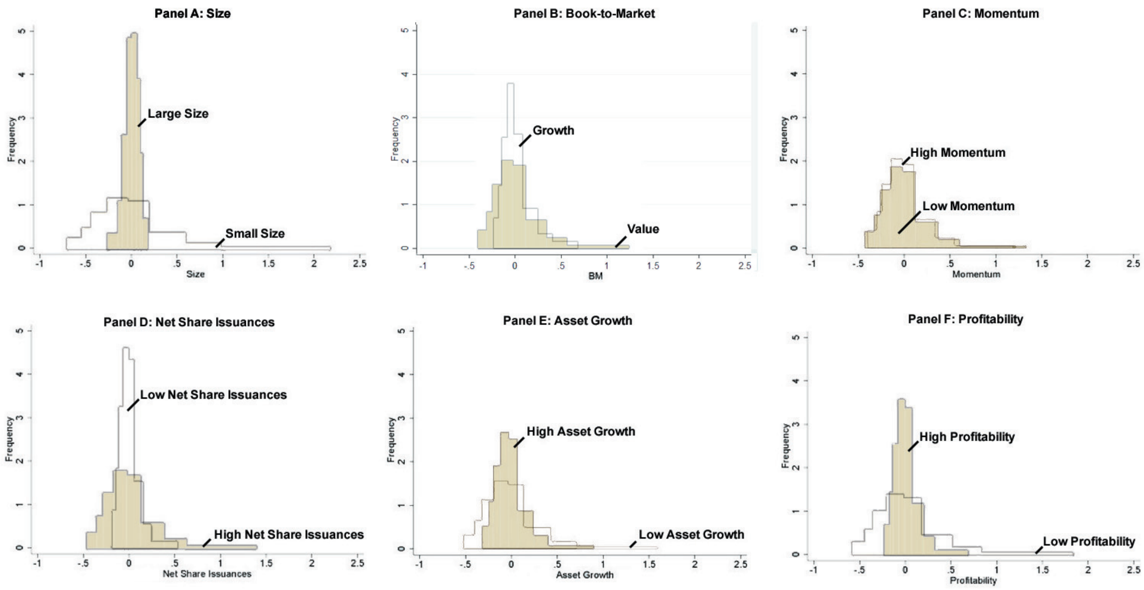

Figure 1. Quantile-Regression Implied Return Densities The figure shows the quantile-regression implied one-month ahead return densities for the *average* stock in extreme firm-fundamental deciles. The firm fundamentals are *Size* (Panel A), *BookToMarket* (Panel B), *Momentum* (Panel C), *ShareIssuances* (Panel D), *AssetGrowth* (Panel E), and *Profitability* (Panel F). The firm fundamental values of the average firm are calculated by averaging firm fundamentals first by month and decile and then by decile alone. We use the quantile regression estimates reported in Table 1 to calculate the fitted quantiles.

to lie below the fitted fifth quantile estimated using only data up to month t if the quantile is well calibrated. Because the densities of small, growth, and illiquid stocks may be harder to forecast than those of others, we run the calibration exercise on the full sample, but also on subsamples containing small, growth, and illiquid stocks. Small and growth stocks are those with a market size and book-to-market ratio in the bottom quartile in month $t-1$, respectively; illiquid stocks are those with an Amihud (2002) illiquidity proxy value in the top quartile in the same month (see Table A.1 for details about the illiquidity proxy).

Table 2 shows that the shorter-ahead return densities of the stocks in the full sample are well calibrated. For example, 1.1% of the full-sample one-month ahead returns are below the first quantile, whereas 99.1% of them are below the 99th quantile. While the densities of small and growth stocks are also well calibrated, the lower tail estimates of illiquid stocks can be slightly off-target. To wit, 12.1% of the one-month ahead returns of illiquid stocks lie below the tenth quantile. Raising the interval over which returns are calculated deteriorates calibration. Most strikingly, only 82.8% of the five-year ahead returns are below the 99th quantile. We speculate that the worse calibration could possibly be an artefact of longer-ahead returns being more non-normally distributed than shorter-ahead returns.

The good calibration of the fitted quantiles of the shorter-ahead returns render us optimistic that the moment-implied quantile-regression skewness forecasts for the same returns are also well calibrated. We turn to the skewness forecasts in the next subsection.

3.2 Comparison of Stock Skewness Forecasts

3.2.1 Descriptive Statistics and Correlations

Table 3 offers descriptive statistics for realized skewness, the quantile-regression based skewness forecasts, and the other skewness forecasts. Since we require option data to calculate realized skewness, Panel A reports descriptive statistics for the subsample of stocks with option contracts written on them. In contrast, Panel B reports descriptive statistics for the full sample of stocks, for all variables mentioned above except realized skewness.

Starting with the subsample of stocks with option contracts, we find that the mean value of *RealizedSkew* is negative, but increases with the length of the return interval (Panel A). For example, the mean realized skewness of the one-month ahead return is –0.69, while the mean realized skewness of the five-year ahead return is –0.06. Consistent with the observation that the realized skewness estimates capture the skewness of *log* returns, while the quantile-regression based forecasts capture the skewness of *discrete* returns, the mean values of *QRSkew* and *NaiveQRSkew* exceed the mean values of *RealizedSkew*. Nonetheless, identical to the realized skewness estimates, they also increase with the length of the return interval, from 0.22 for the one-month ahead return to 2.21 for the five-year ahead return. The only other direct forecast of skewness, *OLSSkew*, also attracts a positive mean value.[11] Looking at the

[11] Because *LogitSkew*, *HistoricalSkew*, and *MaxSkew* are *indirect* skewness forecasts, their mean values are not informative about whether stock returns are positively or negatively skewed.

full sample, we obtain similar descriptive statistics, suggesting that the subsample of stocks with option contracts is representative of the full sample (Panel B).

Table 4 shows the mean cross-sectional correlations between the above variables, calculated using either stocks with option contracts (above diagonal) or all stocks (below diagonal). The table yields the following conclusions: First, the moment-implied quantile-regression based skewness forecasts are always more positively correlated with realized skewness than the other forecasts, possibly hinting at a higher forecasting power. Second, the moment-implied quantile-regression based forecasts are often only weakly correlated with the corresponding naive forecasts (correlations between 0.24 and 0.89). Third, the moment-implied quantile-regression based skewness forecasts are only weakly correlated with *OLSSkew*, *HistoricalSkew*, and *MaxSkew* (correlations between 0.11 and 0.61), but more strongly with *LogitSkew* (correlations close to 0.75). Finally, the correlation between each forecast and realized skewness increases with the interval over which realized skewness is measured. The last observation could be interpreted as suggesting that all forecasts capture the skewness of long-ahead returns. Notwithstanding, we believe a more plausible explanation is that longer-ahead realized skewness is estimated with greater precision than shorter-ahead realized skewness.

3.2.2 Horse Races

We perform two tests to compare the ability of the skewness forecasts to forecast realized skewness. In the first test, we run cross-sectional regressions of *RealizedSkew* on each skewness forecast ("unbiasedness test"). The regression model is given by:

$$RealizedSkew = \alpha + \beta SkewnessForecast + \epsilon, \tag{19}$$

where $RealizedSkew \in \{RealizedSkew_{t+1}, RealizedSkew_{t+1,t+12}, RealizedSkew_{t+1,t+60}\}$, and $SkewnessForecast \in \{QRSkew, NaiveQRSkew, OLSSkew, LogitSkew, HistoricalSkew, MaxSkew\}$. α and β are parameters, and ϵ the residual. Using one of the quantile-regression based forecasts as exogenous variable, we match the return interval used to create the forecast with the return interval underlying *RealizedSkew*. So, for example, we regress *Realized-Skew*$_{t+1,t+60}$ on $QRSkew_{t+1,t+60}$. Because it is unclear over which return interval the other forecasts predict skewness, we regress each *RealizedSkew* variable on them. We expect a less biased forecast to produce a constant closer to zero and a slope coefficient closer to one. In addition, we expect a more efficient forecast to produce a higher R-squared, signalling a higher correlation between realized skewness and the skewness forecast.

Table 5 shows the mean values of the constants and slope coefficients obtained from estimating cross-sectional regression (19) over our sample period. It further shows mean R-squareds and their rank. Panels A, B, and C use $RealizedSkew_{t+1}$, $RealizedSkew_{t+1,t+12}$, and $RealizedSkew_{t+1,t+60}$ as endogenous variables, respectively. Results show that all skewness forecasts are far from being unbiased forecasts of realized skewness. For example, the mean constant and slope coefficient estimates obtained from the regression of $RealizedSkew_{t+1}$ on $QRSkew_{t+1}$ are –0.75 and 0.45, respectively (Panel A). The higher mean levels of $QRSkew$ compared to those of *RealizedSkew* are likely to result from $QRSkew$ capturing the skewness of discrete returns, while *RealizedSkew* captures the skewness of log returns.

Mean R-squareds suggest that the moment-implied quantile-regression based forecasts are the strongest predictors of realized skewness. For example, $QRSkew_{t+1,t+60}$ explains an average of 4.2% of the variations in $RealizedSkew_{t+1,t+60}$, which is 1.1% percentage points higher than the second best-performing forecast (Panel C). In contrast, *NaiveQRSkew* only performs well in forecasting the two longer-ahead realized skewness variables, but not the shortest-ahead one. Of the skewness forecasts used in the prior literature, *LogitSkew* is a strong contender of *QRSkew*. While *LogitSkew* is calculated from one-year ahead returns, its mean R-squared is closest to the mean R-squared of *QRSkew* in models fitting $RealizedSkew_{t+1}$ (0.8% vs. 0.9%; Panel A). In absolute terms, *LogitSkew* also performs well in predicting longer-ahead realized skewness. However, in relative terms, $QRSkew$ significantly outperforms *LogitSkew* in capturing the skewness of one-year ahead returns (R-squareds 2.4% vs. 3.3%; Panel B) or five-year ahead returns (R-squards 3.1% vs. 4.2%; Panel C). Neither *OLSSkew*, *HistoricalSkew*, nor *MaxSkew* are strong predictors of realized skewness over any return interval.

In the second test, we sort stocks into quintile portfolios according to the values of the skewness forecasts in month $t-1$. We then calculate the realized skewness of the portfolios starting from month t. To do so, we first average realized skewness by portfolio and portfolio formation month and then by portfolio alone. We also calculate the difference in realized skewness across the extreme portfolios plus the rank of the difference. Table 8 shows the results, with panels A, B, and C using $RealizedSkew_{t+1}$, $RealizedSkew_{t+1,t+12}$, and $RealizedSkew_{t+1,t+60}$ as realized skewness variable, respectively. Sorting stocks into portfolios according to the quantile-regression based forecasts, we again match the return interval used to create $QRSkew$ with the return interval used to create *RealizedSkew*. In contrast, sorting stocks according to the other forecasts, we use each of the *RealizedSkew* variables.

The results derived from the portfolio formation exercises align with those derived from the unbiasedness regressions. The moment-implied quantile-regression based forecasts always produce the largest spread in realized skewness across the extreme portfolios. Only when forecasting skewness over short horizons can *MaxSkew* sometimes become a strong contender (Panel A). Only when forecasting skewness over long horizons can the naive quantile-regression based forecasts and *LogitSkew* become strong contenders (Panel C).

3.2.3 Optimal Forecast Combinations

As a next step, we form optimal combinations of the skewness forecasts. In doing so, we first transform each cross-section of realized skewness values and skewness forecast values into standard normal variables. We do so because we are interested in the forecasts' ability to correctly rank stocks according to their realized skewness, and not in their ability to exactly fit realized skewness. Also, unless we standardize variables, the indirect forecasts would be at a disadvantage in fitting realized skewness compared to the direct forecasts.

We follow Bates and Granger (1969) and Granger and Ramanathan (1984) in optimally combining forecasts. In particular, we calculate the optimal combination weights from the following cross-sectional non-linear least squares (NLS) regression:

$$RealizedSkew = \frac{1}{1+\sum_k e^{\beta_k}} QRSkew + \sum_{k=1}^{K} \left(\frac{e^{\beta_k}}{1+\sum_k e^{\beta_k}} SkewnessForecast_k \right) + \epsilon, \quad (20)$$

where $RealizedSkew \in \{RealizedSkew_{t+1}, RealizedSkew_{t+1,t+12}, RealizedSkew_{t+1,t+60}\}$, and $QRSkew$ is the moment-implied quantile-regression based skewness forecast calculated from the same return interval as $RealizedSkew$ is. $SkewnessForecast_k \in \{OLSSkew, Logit\text{-}Skew, HistoricalSkew, MaxSkew\}$. β_k is a regression parameter, ϵ the residual, and K the number of other forecasts used in the regression. The optimal weight of the quantile-regression based forecast is $\frac{1}{1+\sum_k e^{\beta_k}}$; the optimal weights of the others are $\frac{e^{\beta_k}}{1+\sum_k e^{\beta_k}}$. The optimal weights are restricted to lie between zero and one and to sum up to one.

Table 7 shows the mean values (Panel A), median values (Panel B), and interquartile ranges (Panel C) of the optimal weights obtained from regression (20). The moment-implied quantile-regression based forecasts are always associated with higher mean (or median) optimal weights than the other skewness forecasts. This result holds independent of whether we let them compete with one other forecast or with all others. For example, when using all skewness forecasts to forecast $RealizedSkew_{t+1,t+60}$, $QRSkew$ is associated with a mean (median) weight of 29% (30%), while no other forecast is associated with a mean (median) weight above 24% (24%). Despite this, all forecasts — not only $QRSkew$ — contribute to correctly ranking stocks according to realized future skewness. For example, in the models including all skewness forecasts, no forecast ever has a mean or median weight below 9%. Also noteworthy is that the optimal weights are relatively stable over the sample period, with the interquartile ranges of the $QRSkew$, $OLSSkew$, and $MaxSkew$ weights never exceeding 10%.

Figure 2 plots the twelve-month moving average weights from the combination regression including $QRSkew_{t+1,t+12}$ and all other forecasts.[12] It confirms that the weights associated

[12]Figures showing the estimates of the other combination regressions produce qualitatively similar conclu-

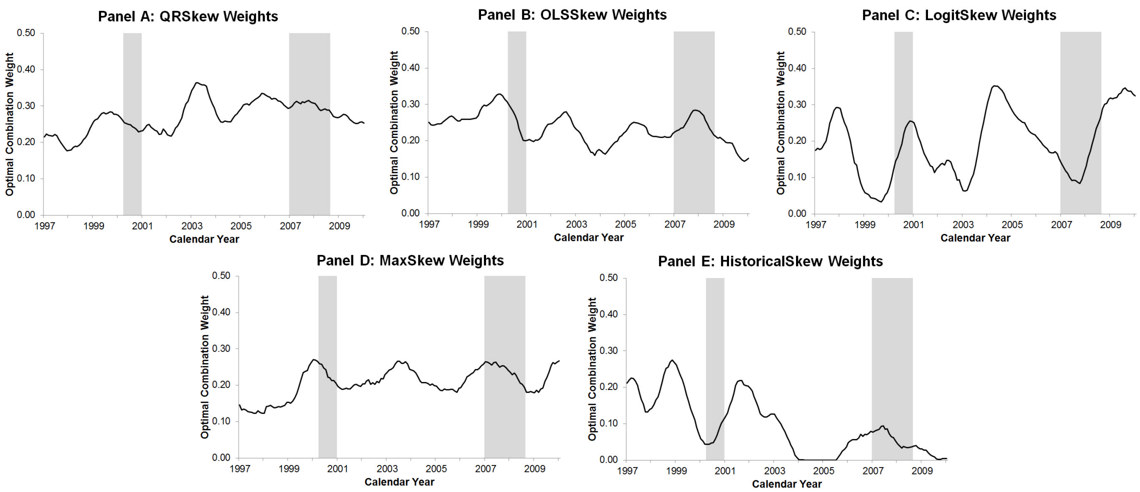

Figure 2. Optimal Combination Weights The figure shows twelve-month moving averages of the optimal combination weights obtained from cross-sectional regression (20), using $RealizedSkew_{t+1,t+12}$ as endogenous variable and $QRSkew_{t+1,t+12}$ (Panel A), $OLSSkew$ (Panel B), $LogitSkew$ (Panel C), $MaxSkew$ (Panel D), and $HistSkew$ (Panel E) as exogenous variables. The regressions are run over all months in our sample period for which we are able to calculate $RealizedSkew_{t+1,t+12}$ over the subsequent twelve months (January 1997 to December 2010). The grey-shaded bars in the figure indicate NBER-defined recession periods.

with the quantile-regression based skewness forecasts do not change much over time. Neither do those associated with $OLSSkew$ and $MaxSkew$. Surprisingly, the weight associated with $LogitSkew$ oscillates around a level of 15%, with little evidence to suggest that the economic state drives the variations. Finally, the weight associated with $MaxSkew$ markedly decreases over time, with it frequently being equal to zero over the 2005–2010 period.

Overall, the tests in this subsection suggest that the moment-implied quantile-regression based skewness forecast are more accurate predictors of realized skewness than the other skewness forecasts. Notwithstanding these conclusions, they also suggest that the other skewness forecasts often possess important incremental information.

sions. In the interest of brevity, we thus omit them.

3.3 The Pricing of Expected Stock Skewness

We now return to the question of whether stock markets price expected skewness. In doing so, we sort stocks into quintile portfolios according to their month $t-1$ values of the skewness forecasts. The first quintile portfolio contains the stocks with a low expected skewness; the fifth contains the stocks with a high expected skewness. We value-weight the portfolios and hold them over month t.[13] We also form a spread portfolio that is long on the highest quintile portfolio and short on the lowest ("5-1"). In addition to mean portfolio returns, we also calculate Hou et al.'s (2014) Q-factor model alpha and Fama and French's (2015) five-factor model alpha for the spread portfolios. We obtain the alphas from time-series regressions of the spread portfolio's return on the appropriate benchmark factors.[14]

Table 8 reports the results of the portfolio formation exercises. Plain numbers are risk premia estimates, while numbers in square parentheses are t-statistics.[15] Neither the moment-implied (Panel A) nor the naive quantile-regression based forecasts (Panel B) are ever statically or economically significantly related to stock returns. For example, $QRSkew_{t+1,t+12}$ produces a risk premium estimate of –0.45% per month (t-stat: –0.84). While the signs of the risk premia estimates obtained from the two shorter-ahead forecasts are negative and thus consistent with prior empirical work, the longest-ahead forecasts always produce positive estimates. This result is noteworthy because long-term investors are expected to care most about skewness over long-term (and not short-term) future horizons. Correcting for the Q-theory or Fama and French (2015) factors does not materially change our conclusions.

Turning to the skewness forecasts used in the prior literature (Panel C), our results confirm those in Boyer et al. (2010). To wit, $OLSSkew$ produces a significantly negative risk premium

[13] Using equal weights instead of value weights does not change our conclusions. Thus, we do not report the results from the equally-weighted portfolios, but make them available upon request.

[14] We thank Lu Zhang for sending us their benchmark factor data (the excess market return and the returns of a size spread portfolio, an investment spread portfolio, and a profitability spread portfolio). We thank Kenneth French for making their benchmark factor data (the excess market return and the returns of a size spread portfolio, a book-to-market spread portfolio, an investment spread portfolio, and a profitability spread portfolio) available on his website (*mba.tuck.dartmouth.edu/pages/faculty/ken.french/*).

[15] We always calculate t-statistics using Newey and West's (1987) formula with a lag length (l) of twelve.

of –0.70% per month (t-stat: –2.02). However, deviating from Bali et al.'s (2011) and Conrad et al.'s (2014) results, *LogitSkew* and *MaxSkew* do not produce significantly negative risk premia. Our risk premium estimate for *LogitSkew* differs from the one reported in Conrad et al. (2014) because we study a more recent sample period. Using their sample period (1972–2009), we find that *LogitSkew* has a significantly negative premium of –0.73 (t-stat: –2.24). Our risk premium estimate for *MaxSkew* differs from the one reported in Bali et al. (2011) because we value-weight stocks, while they equally-weight them. Using equal weights, we find that *MaxSkew* has a moderately significantly negative premium of –0.82 (t-stat: –1.82).

Table 9 shows the results from portfolio formation exercises, using the optimal forecast combinations as sorting variable. The column labels indicate the quantile-regression based forecast used in the combination; the rows labels indicate the other forecast(s). Except in the row labeled "all (equal-weights)," we create the combination forecasts using weights obtained from cross-sectional regression (20), using standardized realized skewness over months $t-60$ to $t-1$ as endogenous variable and the standardized skewness forecasts at the end of month $t-61$ as exogenous variables. We combine the weights with the skewness forecast values in month $t-1$. The combination forecasts in the row labeled "all (equal-weights)" are equally-weighted averages of the (standardized) skewness forecasts. The table shows that no skewness forecast combination ever produces a significant risk premium. This conclusion holds independent of the realized skewness variable used to construct the skewness forecast combination or whether or not we adjust for the Q-theory or Fama-French five-factor model factors.

Table 10 shows the results from Fama-MacBeth (1973) regressions of stock returns on the skewness forecasts and control variables. Bold numbers are monthly risk premia estimates, whereas numbers in parentheses are t-statistics. As control variables, we use the *MarketBeta*, *Size*, *BookToMarket*, *Momentum*, *AssetGrowth*, *Profitability*, and *Volatility*.[16] Supporting our former findings, the quantile-regression based skewness forecasts never produce significant risk premia. However, in contrast to before, *OLSSkew* also no longer produces a significantly

[16]We describe the construction of the controls in Table A.1.

negative risk premium, while *MaxSkew* does. Given that Fama-MacBeth (1973) regressions equally-weight observations, it appears that the significantly negative risk premium of *OLSSkew* reported above is largely driven by larger stocks, while the significantly negative risk premium of *MaxSkew* reported in the current tests is mostly driven by smaller stocks.[17] The control variables produce slope coefficient estimates consistent with the prior literature.

Table 11 repeats the Fama-MacBeth (1973) regressions, using the skewness forecast combinations as exogenous variables. Consistent with Table 9, the column labels indicate the quantile-regression based forecast used in the combination; the rows indicate the other forecast(s). To save space, the table only reports the risk premium estimates of the combinations. In accordance with our former results, no combination ever produces a significant risk premium. While the pattern is not significant, risk premia estimates again increase with the length of the return interval over which expected skewness is calculated.

Overall, the tests in this subsection produce little evidence suggesting that stock markets negatively price expected stock skewness.

4 Conclusion

We propose an efficient and unbiased estimator of the skewness of a stock's return. The estimator can be used to forecast the skewness of stock returns calculated over any conceivable return interval. To implement the estimator, we perform quantile regressions fitting stocks' future return quantiles. Using the fitted quantiles, we calculate forecasts of a stock's future return skewness. We use Neuberger's (2012) realized skewness to benchmark the quantile-regression based skewness forecasts with other popular skewness forecasts used in the literature. Results suggest that the quantile-regression based skewness forecasts strongly dominates the other skewness forecasts, with the possible exception of Conrad et al.'s (2014) logit-model based forecast. Finally, we use the quantile-regression based skewness forecasts and various

[17]Unreported equally-weighted portfolio formation exercises confirm this claim (available upon request).

combinations of alternative skewness forecasts as pricing factors in asset pricing tests. Results fail to suggest that stock markets price expected skewness, casting doubt on a growing empirical literature suggesting that expected stock skewness is negatively priced.

References

Amaya, D., Christoffersen, P., Jacobs, K., Vasquez, A., 2015. "Does realized skewness predict the cross-section of equity returns? " *Journal of Financial Economics* 118, 135-167.

Amihud, Y., 2002. "Illiquidity and stock returns: Cross-section and time-series effects." *Journal of Financial Markets* 5, 31-56.

Andersen, T. G., Bollerslev, T., 1998. "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts." *International Economic Review* 39, 885-905.

Angrist, J. D., J.-s. Pischke. 2009. "Mostly harmless econometrics: An empiricists companion." *Princeton University Press, Princeton.*

Bakshi, G., Kapadia, N., Madan, D., 2003. "Stock return characteristics, skew laws, and the differential pricing of individual equity options." *Review of Financial Studies* 16, 101-143.

Bali, T. G., Cakici, N., Whitelaw, R. F., 2011. "Maxing out: Stocks as lotteries and the cross-section of expected returns." *Journal of Financial Economics* 99, 427-446.

Barberis, N., Huang, M., 2008. "Stocks as lotteries: The implications of probability weighting for security prices." *American Economic Review* 98, 2066-2100.

Bates, J. M., Granger, C. W. J., 1969. "The combination of forecasts." *Operations Research Quarterly* 20, 451-468.

Boyer, B., Mitton, T., Vorkink, K., 2010. "Expected idiosyncratic skewness." *Review of Financial Studies* 23, 169-202.

Boyer, B., Vorkink, K., 2014. "Stock options as lotteries." *Journal of Finance* 69, 1485-1528.

Brunnermeier, M. K., Gollier,C., Parker, J. A., 2007. "Optimal beliefs, asset prices, and the preference for skewed returns." *American Economic Review* 97, 159-165.

Campbell, J. Y., Hilscher, J., Szilagyi, J., 2008. "In search of distress risk." *The Journal of Finance* 63, 2899-2939.

Carr, P., Wu, L., 2009. "Variance risk premiums." *Review of Financial Studies* 22, 1311-1341.

Carr, P., Wu, L., 2007. "Stochastic skew in currency options." *Journal of Financial Economics* 86, 213-247.

Cass, D., Stiglitz, J., 1970. "The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds." *Journal of Economic Theory* 2, 122-160.

Cenesizoglu, T., Timmermann, A., 2008. "Is the distribution of stock returns predictable?" *Working Paper, University of California San Diego.*

Chen, J., Hong, H., Stein, J., 2001. "Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. " *Journal of Financial Economics* 61, 345-381.

Conine, T. E., Tamarkin, M. J., 1981. "On diversification given asymmetry in returns." *The Journal of Finance* 36, 1143-1155.

Conrad, J., Dittmar, R. F., Ghysels, E., 2013. "Ex ante skewness and expected stock returns." *The Journal of Finance* 68, 85-124.

Conrad, J., Kapadia, N., Xing, Y., 2014. "Death and jackpot: Why do individual investors hold overpriced stocks?" *Journal of Financial Economics* 113, 455-475.

Dimson, E., 1979. "Risk measurement when shares are subject to infrequent trading." *Journal of Financial Economics* 7, 197-226.

Fama, E. F., French, K. R., 1992. "The cross-section of expected stock returns." *The Journal of Finance* 47, 427-465.

Fama, E. F., French, K. R., 2008. "Dissecting anomalies." *The Journal of Finance* 63, 1653-1678.

Fama, E. F., French, K. R., 2015. "A five factor asset pricing model." *Journal of Financial Economics*, 116, 1-22.

Fama, E. F., MacBeth, J. D., 1973. "Risk, return, and equilibrium: Empirical tests." *Journal of Political Economy* 71, 607-636.

Granger, C., Ramanathan, R., 1984. "Improved methods of combining forecasts." *Journal of Forecasting* 3, 197-204.

Harvey, C. R., Siddique, A., 1999. "Autoregressive conditional skewness." *Journal of Financial and Quantitative Analysis* 34, 465-488.

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

Hong, H., Stein, J., 2003. "Differences of opinion, short-sales constraints and market crashes." *Review of Financial Studies* 16, 487-525.

Hou, K., Chen, X., Zhang, L., 2014. "Digesting anomalies: An investment approach." *Review of Financial Studies* 28, 650-705.

Kim, T.-h., White, H., 2004. "On more robust estimation of skewness and kurtosis." *Finance Research Letters* 1, 56-73.

Koenker, R., 2005. "Quantile Regressions." *Cambridge University Press, New York.*

Koenker, R., Bassett, G., 1978. "Regression quantiles." *Econometrica* 50, 43-61.

Konstantinidi, T., Pope, P. F., 2015. "Forecasting risk in earnings. "}Contemporary Accounting Research*, doi: 10.1111/1911-3846.12158.*

Lewellen, J., Nagel, S., 2006. "The conditional CAPM does not explain asset pricing anomalies." Journal of Financial Economics *82, 289-314.*

Mitton, T., Vorkink, K., 2007. "Equilibrium underdiversification and the preference for skewness." Review of Financial Studies *20, 1255-1288.*

Newey, W. K., West, K. D., 1987. "A simple positive-definite heteroscedasticity and autocorrelation consistent covariance matrix." Econometrica *55, 703-708.*

Neuberger, A., 2012. "Realized skewness." Review of Financial Studies *25, 3423-3455.*

Rehman, Z., Vilkov, G., 2012. " Risk-neutral skewness: Return predictability and its sources." *Working Paper, Goethe University.*

Rubinstein, M., 1973. The fundamental theorem of parameter-preference security valuation." *Journal of Financial and Quantitative Analysis* 8, 61-69.

Rubinstein, M., 1976. "The valuation of uncertain income streams and the pricing of options." *Bell Journal of Economics* 7, 407-425.

Scott, R. C., Horvath, P. A., 1980. "On the direction of preference for moments of higher order than the variance." *The Journal of Finance* 35, 915-919.

Simkowitz, M. A., Beedles, W. L., 1978. "Diversification in a three-moment world. " *The Journal of Financial and Quantitative Analysis* 13, 927-941.

Stilger, P. S., Kostakis, A., Poon, S.-h., 2014. "What does risk-neutral skewness tell us about future stock returns?," *Working Paper, Manchester Business School.*

Xing, Y., Zhang, X., Zhao, R., 2010. "What does the individual option volatility smirk tell us about future equity returns?" *Journal of Financial and Quantitative Analysis* 45, 641-662.

Table A.1
Variable Construction

The table shows information on the analysis variables. The first column gives the variables' mnemonics. The second column gives their full names. In the third column, we outline how we calculate the variables. Panel A contains the skewness forecasts and the realized skewness proxy, Panel B the variables used to construct the quantile regression-based skewness forecast, and Panel C other analysis variables. Where applicable, we show the COMPUSTAT data item number in parentheses.

Mnemonic	Name	Description
Panel A: Skewness proxies		
RealizedSkew	Neuberger (2012) Realized Skewness	Realized skewness calculated from daily stock return and option data over various future periods.
QRSkew/ *NaiveQRSkew*	Quantile Regression-Based Skewness Forecasts	Skewness forecasts derived from quantile regressions of future returns on lagged firm characteristics.
OLSSkew	Boyer et al. (2010) OLS Skewness Forecast	Skewness forecast derived from cross-sectional OLS regressions of realized future skewness calculated from daily returns over the next 60 months on lagged firm characteristics.
LogitSkew	Conrad et al. (2014) Logit Skewness Forecast	Skewness forecast derived from logit models of a dummy variable equal to one if the return over the next twelve months exceeds 100
MaxSkew	Bali et al. (2014) Skewness Forecast	Maximum daily return over the prior one month.
HistoricalSkew	Historical Skewness	Realized skewness calculated from daily return data over the past 60 months.
Panel B: Firm characteristics		
Hist.Volatility	Historical volatility	Stock return volatility calculated using daily return data over the past 60 months.
Hist.Skew	Historical skewness	Stock return skewness calculated from daily return data over the past 60 months.
Size	Market capitalization	Log of share price times number of common shares outstanding.
BookToMarket	Book-to-market ratio	Log of the ratio of book value of equity at the end of fiscal year t-1 to the market value of equity at the same time: the book value of equity is total assets (6) for year t-1, minus liabilities (181), plus balance sheet deferred taxes and investment tax credit (35), minus preferred stock liquidating value (10) or redemption value (56), or carrying value (130). Market equity is given by *Size*.
Momentum	Stock Return Momentum	Log of gross return continuously compounded over months t-12 to t-2.
ShareIssue	Net stock issues	Log of split-adjusted shares outstanding at fiscal year end in t-1 divided by split adjusted shares outstanding at fiscal year end in t-2; the split adjusted shares outstanding is shares outstanding (25) times the adjustment factor (27).
Accruals	Accruals	Change in operating working capital per split-adjusted share from t-2 to t-1, divided by book equity per split-adjusted share at t-1. Operating working capital is current assets (4) minus cash and short-term investments (1) minus current liabilities (5) plus debt in current liabilities (34).
AssetGrowth	Asset growth	Log of the ratio of assets per split-adjusted share at fiscal year end in t-1 divided by assets per split-adjusted share at fiscal year end in t-2.
Profitability	Profitability	Equity income in t-1 (18) minus dividends on preferred stock in t-1 (19) plus deferred taxes in t-1 (50) divided by book value of equity at t-1.
Turnover	Share turnover	Average daily turnover over the most recent month.
NASDAQ	NASDAQ dummy variable	Dummy variable equal to one if a share trades on the NASDAQ, else zero.
Panel C: Additional variables		
MarketBeta	Conditional market beta	Market beta estimated using a regression of the stock return on the excess market return and several lagged market returns run over daily data over the prior twelve months (see Lewellen and Nagel (2004)).
Volatility	Historical volatility	Historical volatility calculated from daily return data over the past three months.
ShareIlliquidity	Share illiquidity proxy	Ratio of absolute daily return to daily trading volume averaged over the past twelve months (see Amihud (2002)).

Table 1
Quantile Regression Estimates
The table shows the results from the following quantile regressions:

$$Q_{R^i_{t,t+1}}(\tau|\mathbf{X}_t) = \alpha_k + \mathbf{X}_t\beta + F_u^{-1}(\tau),$$

where $Q_Y(\tau|\mathbf{X})$ is the τth quantile of Y conditional on $\mathbf{X}$, $R^i_{t,t+1}$ is stock i's return from the end of month t to the end of month $t+1$, and $\mathbf{X}_t$ is a vector of variables measured at the end of month t. The $\mathbf{X}_t$ variables contain: *Hist. Vol*, *Hist. Skew*, *Size*, *BookToMarket*, *Momentum*, *ShareIssue*, *Accruals*, *AssetGrowth*, *Profitability*, *Turnover*, *NASDAQ*, and a constant (*Cons*). We describe their construction in Table A.1. α_k is an industry fixed-effect and β a vector of free parameters. The industry fixed-effects are based on Kenneth French's 17 industry classification scheme. We perform the quantile regressions over rolling windows featuring the previous 20 years of monthly data. The first window ends in December 1987; the last window ends in December 2010. The estimation windows are rolled forward on an annual basis. In the first column, we show the fitted quartile (τ). The next columns show the mean slope coefficients of the variables in $\mathbf{X}_t$ (in bold) and the fraction of months in which the slope coefficient is significant at the 90% confidence level or better (in parentheses). The final column shows the mean R-squared ($\bar{R}^2$). The sample period is January 1968 to December 2010.

Quantile	*Hist. Volatility*	*Hist. Skew*	*Size*	*BookTo Market*	*Momentum*	*Share Issue*	*Accruals*	*Asset Growth*	*Profi- tability*	*Turn- over*	*NAS- DAQ*	*Cons.*	$\bar{R}^2$
1 (1st Quantile)	**-3.58**	**0.28**	**1.06**	**2.07**	**0.01**	**-7.48**	**-0.11**	**-3.16**	**4.63**	**-29.22**	**-1.18**	**-28.83**	0.132
	1.00	0.93	1.00	1.00	1.00	1.00	0.18	1.00	1.00	1.00	0.64	1.00	
5	**-2.80**	**0.20**	**0.61**	**1.56**	**0.01**	**-5.04**	**-0.18**	**-1.55**	**2.62**	**-15.55**	**-0.59**	**-15.16**	0.104
	1.00	0.96	1.00	1.00	1.00	1.00	0.93	1.00	1.00	1.00	0.75	1.00	
10	**-2.32**	**0.17**	**0.44**	**1.31**	**0.01**	**-4.54**	**-0.17**	**-1.17**	**1.94**	**-10.66**	**-0.24**	**-10.30**	0.082
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.64	1.00	
25 (1st Quartile)	**-1.44**	**0.13**	**0.25**	**0.98**	**0.01**	**-3.40**	**-0.16**	**-0.66**	**1.20**	**-5.18**	**0.14**	**-4.49**	0.038
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.68	1.00	
50 (Median)	**-0.21**	**-0.02**	**0.14**	**0.48**	**0.00**	**-2.02**	**-0.09**	**-0.23**	**0.57**	**-0.26**	**0.03**	**-0.47**	0.003
	1.00	0.54	1.00	1.00	1.00	1.00	1.00	0.71	1.00	0.57	0.50	0.75	
75 (3rd Quartile)	**0.92**	**-0.29**	**-0.08**	**0.23**	**0.00**	**-0.95**	**-0.06**	**0.62**	**0.06**	**7.55**	**-0.38**	**5.88**	0.009
	1.00	1.00	0.64	0.82	0.71	1.00	0.57	0.61	0.39	1.00	0.75	1.00	
90	**2.89**	**-0.31**	**-0.74**	**0.01**	**-0.01**	**-0.43**	**-0.01**	**0.77**	**-1.16**	**12.12**	**-0.23**	**16.75**	0.053
	1.00	1.00	1.00	0.43	1.00	0.50	0.21	0.57	0.93	1.00	0.75	1.00	
95	**4.71**	**-0.20**	**-1.25**	**-0.04**	**-0.02**	**-1.30**	**0.03**	**0.48**	**-2.70**	**14.12**	**-0.01**	**24.45**	0.096
	1.00	0.71	1.00	0.54	1.00	0.79	0.46	0.46	1.00	1.00	0.57	1.00	
99 (99th Quantile)	**10.07**	**0.17**	**-2.57**	**-0.19**	**-0.04**	**-2.62**	**0.21**	**-0.82**	**-8.43**	**20.06**	**0.23**	**42.59**	0.192
	1.00	0.57	1.00	0.54	1.00	0.68	0.79	0.61	1.00	1.00	0.46	1.00	

Table 2
Unbiasedness Tests of the Fitted Quantiles
The table shows the proportions of future returns that lie below quantile-regression based estimates of their quantiles. The quantile regressions are performed over rolling windows featuring the previous 20 years of monthly data. The first window ends in December 1987, while the last window ends in December 2010. The estimation windows are rolled forward on an annual basis. We fit the first (Pct1), fifth (Pct5), tenth (Pct10), 25th (Pct25), 50th (Pct50), 75th (Pct75), 90th (Pct90), 95th (Pct95), and 99th (Pct99) quantiles. To create the fitted quantiles, we combine the quantile-regression estimates from the rolling window ending in December of the previous year with the values of the explanatory variables in month t. We then compare the fitted quantiles of the return over month $t+1$ ("monthly return;" Panel A), the return over months $t+1$ to $t+12$ ("annual return;" Panel B), and the return over months $t+1$ to $t+60$ ("five-year return;" Panel C) with their respective return realizations. We do the comparison separately for the full sample, small stocks, illiquid stocks, and growth stocks. Small and growth stocks are those with a market capitalization or book-to-market ratio in the bottom quartile in month t; illiquid stocks are those with an Amihud (2002) illiquidity proxy value in the top quartile in the same month. We estimate the quantile regressions over the January 1968-December 2010 period, but we run the unbiasedness tests over the December 1987-December 2010 period.

	Proportion of Future Returns Below								
	Pct1	Pct5	Pct10	Q1	Median	Q3	Pct90	Pct95	Pct99
Panel A: One Month-Ahead Returns									
All Stocks	0.011	0.051	0.100	0.244	0.498	0.754	0.905	0.953	0.991
Small Stocks	0.011	0.052	0.101	0.245	0.503	0.759	0.903	0.951	0.989
Illiquid Stocks	0.012	0.064	0.121	0.275	0.512	0.763	0.902	0.948	0.987
Growth Stocks	0.014	0.058	0.108	0.254	0.508	0.756	0.904	0.952	0.990
Panel B: Twelve Month-Ahead Returns									
All Stocks	0.018	0.064	0.112	0.243	0.475	0.720	0.869	0.919	0.961
Small Stocks	0.015	0.060	0.110	0.248	0.487	0.732	0.872	0.919	0.960
Illiquid Stocks	0.016	0.061	0.108	0.238	0.475	0.728	0.884	0.937	0.982
Growth Stocks	0.027	0.085	0.139	0.277	0.508	0.734	0.869	0.917	0.960
Panel C: Five Year-Ahead Returns									
All Stocks	0.021	0.056	0.101	0.229	0.430	0.633	0.752	0.793	0.828
Small Stocks	0.017	0.038	0.076	0.208	0.421	0.641	0.761	0.800	0.835
Illiquid Stocks	0.020	0.059	0.108	0.244	0.449	0.642	0.758	0.804	0.843
Growth Stocks	0.025	0.076	0.133	0.280	0.489	0.671	0.772	0.805	0.833

Table 3
Descriptive Statistics for Skewness Forecasts and Realized Skewness
The table reports descriptive statistics for the various stock return skewness forecasts and the realized stock return skewness measure. The descriptive statistics contain the mean (Mean), the standard deviation (StD), and the first (Pct1), tenth (Pct10), 25th (Pct25), 50th (Pct50), 75th (Pct75), 90th (Pct90), and 99th percentiles (Pct99). Panel A shows descriptive statistics for the subsample of stocks with option contracts written on them; Panel B shows descriptive statistics for the full sample. *RealizedSkew* is Neuberger's (2012) realized skewness measure; its subscript indicates the length of the period over which the stock return underlying the measure is based. *QRSkew* and *NaiveQRSkew* are quantile-regression based skewness forecasts; their subscripts indicate the length of the period over which the stock return underlying the measure is based. *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. See Table A.1 for details. The sample period is December 1987 to December 2010.

	Mean	StD	Pct1	Pct10	Q1	Median	Q3	Pct90	Pct99
Panel A: Stocks With Option Contracts Written on Them									
Realized Skewness									
$RealizedSkew_{t+1}$	-0.693	3.507	-12.318	-4.225	-1.916	-0.403	0.868	2.617	7.884
$RealizedSkew_{t+1,t+12}$	-0.133	0.264	-0.929	-0.458	-0.273	-0.111	0.030	0.170	0.456
$RealizedSkew_{t+1,t+60}$	-0.063	0.077	-0.302	-0.153	-0.101	-0.055	-0.015	0.022	0.097
Quantile Regression-Based Skewness Forecasts									
$QRSkew_{t+1}$	0.220	0.339	-0.616	-0.221	0.004	0.232	0.451	0.643	0.957
$QRSkew_{t+1,t+12}$	1.021	0.635	-0.673	0.192	0.629	1.061	1.452	1.817	2.299
$QRSkew_{t+1,t+60}$	2.214	0.982	-0.550	0.915	1.690	2.308	2.847	3.351	4.312
Naive Quantile Regression-Based Skewness Forecasts									
$NaiveQRSkew_{t+1}$	0.041	0.033	-0.046	0.003	0.022	0.041	0.059	0.078	0.127
$NaiveQRSkew_{t+1,t+12}$	0.095	0.067	-0.092	0.007	0.054	0.101	0.145	0.177	0.219
$NaiveQRSkew_{t+1,t+60}$	0.020	0.060	-0.167	-0.063	-0.013	0.030	0.061	0.084	0.129
Other Skewness Forecasts									
OLSSkew	0.645	0.397	-0.222	0.120	0.375	0.662	0.888	1.123	1.676
LogitSkew	-4.971	0.797	-7.001	-6.162	-5.422	-4.840	-4.412	-4.080	-3.446
HistoricalSkew	0.241	0.801	-1.630	-0.730	-0.207	0.198	0.630	1.231	2.530
MaxSkew	0.061	0.048	0.012	0.022	0.031	0.047	0.075	0.116	0.248
Panel B: All Stocks									
Quantile Regression-Based Skewness Forecasts									
$QRSkew_{t+1}$	0.531	0.403	-0.448	-0.004	0.252	0.557	0.827	1.028	1.361
$QRSkew_{t+1,t+12}$	1.306	0.612	-0.349	0.518	0.935	1.334	1.718	2.064	2.606
$QRSkew_{t+1,t+60}$	2.348	0.857	-0.214	1.241	1.884	2.467	2.907	3.281	4.145
Naive Quantile Regression-Based Skewness Forecasts									
$NaiveQRSkew_{t+1}$	0.026	0.047	-0.107	-0.035	0.001	0.029	0.054	0.080	0.139
$NaiveQRSkew_{t+1,t+12}$	0.139	0.062	-0.052	0.057	0.103	0.148	0.183	0.210	0.248
$NaiveQRSkew_{t+1,t+60}$	0.060	0.054	-0.104	-0.007	0.032	0.067	0.097	0.121	0.161
Other Skewness Forecasts									
OLSSkew	1.084	0.588	-0.126	0.355	0.663	1.035	1.504	1.872	2.514
LogitSkew	-4.273	1.012	-6.844	-5.610	-4.878	-4.225	-3.621	-3.034	-1.917
HistoricalSkew	0.410	0.901	-1.848	-0.588	-0.088	0.326	0.845	1.585	2.987
MaxSkew	0.087	0.084	0.000	0.022	0.036	0.061	0.106	0.179	0.444

Table 4

Correlations Between Skewness Forecasts and Realized Skewness

The table shows the correlations between the stock return skewness forecasts and realized skewness. We calculate correlations first by month and then average over months. The correlations above the diagonal are calculated from the sample of stocks with option contracts written on them; the correlation below the diagonal are calculated from the full sample. *RealizedSkew* is Neuberger's (2012) realized skewness measure. The variable's subscript indicates the length of the period over which the return underlying the realized skewness variable is calculated. *QRSkew* and *NaiveQRSkew* are skewness forecasts derived from quantile regressions. The variables' subscripts indicate the length of the period over which the return underlying the skewness forecast is calculated. *OLSSkew* is an OLS-based forecast of the realized skewness of the daily return over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. See Table A.1 for more details. The sample period is December 1987 to December 2010.

	RealizedSkew			*QRSkew*			*NaiveQRSkew*			*OLS*	*Logit*	*Hist.*	*Max*
	$t{+}1$	$t{+}1,t{+}12$	$t{+}1,t{+}60$	$t{+}1$	$t{+}1,t{+}12$	$t{+}1,t{+}60$	$t{+}1$	$t{+}1,t{+}12$	$t{+}1,t{+}60$	*Skew*	*Skew*	*Skew*	*Skew*
$RealizedSkew_{t+1}$	1.000												
$RealizedSkew_{t+1,t+12}$	0.153	1.000											
$RealizedSkew_{t+1,t+60}$	0.039	0.291	1.000										
$QRSkew_{t+1}$	0.033	0.142	0.215	1.000	0.859	0.467	-0.272	0.782	0.627	0.607	0.817	0.159	0.416
$QRSkew_{t+1,t+12}$	0.040	0.131	0.224	0.896	1.000	0.721	-0.271	0.890	0.565	0.484	0.752	0.153	0.452
$QRSkew_{t+1,t+60}$	0.045	0.109	0.193	0.674	0.861	1.000	-0.271	0.712	0.237	0.218	0.539	0.109	0.380
$NaiveQRSkew_{t+1}$	0.003	-0.070	-0.061	0.057	0.147	0.233	1.000	-0.258	0.160	-0.165	-0.244	-0.040	-0.272
$NaiveQRSkew_{t+1,t+12}$	0.042	0.116	0.190	0.782	0.889	0.790	0.250	1.000	0.592	0.479	0.743	0.151	0.434
$NaiveQRSkew_{t+1,t+60}$	0.029	0.084	0.135	0.663	0.625	0.477	0.284	0.715	1.000	0.385	0.530	0.091	0.205
OLSSkew	0.013	0.108	0.164	0.435	0.348	0.196	-0.106	0.327	0.278	1.000	0.565	0.102	0.277
LogitSkew	0.031	0.130	0.162	0.686	0.717	0.699	0.069	0.681	0.452	0.319	1.000	0.189	0.546
HistSkew	-0.021	0.006	0.025	0.292	0.230	0.117	-0.178	0.140	0.059	0.239	0.115	1.000	0.206
MaxSkew	0.032	0.128	0.062	0.368	0.385	0.374	0.081	0.385	0.228	0.225	0.411	0.280	1.000

Table 5
Cross-Sectional Regressions of Realized Skewness on Skewness Forecasts
This table shows the results from the following cross-sectional regression:

$$RealizedSkew = \alpha + \beta SkewnessForecast + \epsilon,$$

where $RealizedSkew$ measures the realized stock return skew over the next month ($RealizedSkew_{t+1}$; Panel A), the next twelve months ($RealizedSkew_{t+1,t+12}$; Panel B), or the next five years ($RealizedSkew_{t+1,t+60}$; Panel C); and $SkewnessForecast \in (QRSkew, NaiveQRSkew, OLSSkew,\ LogitSkew,\ HistoricalSkew,\ MaxSkew)$. *QRSkew* and *NaiveQRSkew* are quantile-regression based skewness forecasts based on the same return frequency as the *RealizedSkew* variable with which they are associated; *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. See Table A.1 for details. α and β are free parameters, and ϵ is the residual. We run the regressions separately per month. We report the mean coefficients (in bold), their associated t-statistics (in square parentheses), the average R-squared ($\bar{R}^2$), and the average R-squared's rank ($\bar{R}^2$-Rank). The sample period is December 1987 to December 2010.

	α		β			
	mean est.	t-stat	mean est.	t-stat	$\bar{R}^2$	$\bar{R}^2$-Rank
Panel A: One-Month Ahead Realized Skewness						
$QRSkew_{t+1}$	**-0.76**	[-9.51]	**0.45**	[7.44]	0.009	1
$NaiveQRSkew_{t+1}$	**-0.66**	[-8.78]	**0.91**	[1.53]	0.003	5
OLSSkew	**-0.81**	[-9.44]	**0.22**	[4.75]	0.003	4
LogitSkew	**0.16**	[1.40]	**0.16**	[6.69]	0.008	2
HistoricalSkew	**-0.64**	[-8.83]	**-0.10**	[-7.19]	0.002	6
MaxSkew	**-0.75**	[-9.25]	**1.44**	[3.65]	0.004	3
Panel B: Twelve-Month Ahead Realized Skewness						
$QRSkew_{t+1,t+12}$	**-0.19**	[-20.90]	**0.07**	[17.23]	0.033	1
$NaiveQRSkew_{t+1,t+12}$	**-0.19**	[-21.29]	**0.51**	[16.78]	0.026	2
OLSSkew	**-0.16**	[-20.52]	**0.04**	[8.62]	0.005	5
LogitSkew	**0.07**	[7.89]	**0.04**	[17.68]	0.024	3
HistoricalSkew	**-0.13**	[-22.52]	**0.00**	[1.89]	0.002	6
MaxSkew	**-0.16**	[-21.59]	**0.55**	[13.94]	0.013	4
Panel C: Five-Year Ahead Realized Skewness						
$QRSkew_{t+1,t+60}$	**-0.09**	[-43.71]	**0.01**	[33.74]	0.042	1
$NaiveQRSkew_{t+1,t+60}$	**-0.07**	[-33.97]	**0.29**	[27.36]	0.029	3
OLSSkew	**-0.07**	[-29.89]	**0.01**	[5.38]	0.007	5
LogitSkew	**0.02**	[10.34]	**0.02**	[35.50]	0.031	2
HistoricalSkew	**-0.06**	[-45.88]	**0.00**	[6.89]	0.002	6
MaxSkew	**-0.07**	[-44.67]	**0.24**	[24.22]	0.018	4

Table 6
Average Realized Skewness By Skewness Forecast Portfolio
The table reports the realized skewness of portfolios sorted according to various skewness forecasts. Panels A to C look at the realized skewness of the stock return measured over the next month ($RealizedSkew_{t+1}$; Panel A), the next twelve months ($RealizedSkew_{t+1,t+12}$; Panel B), or the next five years ($RealizedSkew_{t+1,t+60}$, Panel C), respectively. The skewness forecasts are: $QRSkew$, $NaiveQRSkew$, $OLSSkew$, $LogitSkew$, $HistoricalSkew$, and $MaxSkew$. *QRSkew* and *NaiveQRSkew* are quantile-regression based skewness forecasts based on the same return frequency as the *RealizedSkew* variable with which they are associated; *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. See Table A.1 for more details. We form the portfolios as follows: At the end of every month t in our sample period, we sort stocks into portfolios according to the quintile breakpoints of one of the skewness forecasts. Portfolio 1 (Low) contains stocks with low skewness forecast values; portfolio 5 contains stocks with high skewness forecast values. We calculate mean realized skewness by portfolio and month and then by portfolio alone. "5–1" is the spread in mean realized skewness between portfolio 5 and 1. "Accuracy Rank" is the rank of the spread. The sample period is December 1987 to December 2010.

	Skewness Forecast Deciles						Accuracy
Skewness Forecast	1 (Low)	2	3	4	5 (High)	5–1	Rank
Panel A: One-Month Ahead Realized Skewness							
$QRSkew_{t+1}$	-0.835	-0.759	-0.613	-0.542	-0.431	0.404	1
$NaiveQRSkew_{t+1}$	-0.666	-0.615	-0.613	-0.650	-0.638	0.028	5
OLSSkew	-0.710	-0.693	-0.692	-0.716	-0.513	0.197	4
LogitSkew	-0.804	-0.721	-0.695	-0.599	-0.429	0.375	2
HistoricalSkew	-0.590	-0.645	-0.663	-0.630	-0.777	-0.188	6
MaxSkew	-0.894	-0.706	-0.627	-0.536	-0.541	0.353	3
Panel B: Twelve-Month Ahead Realized Skewness							
$QRSkew_{t+1,t+12}$	-0.175	-0.146	-0.131	-0.103	-0.066	0.109	1
$NaiveQRSkew_{t+1,t+12}$	-0.176	-0.143	-0.127	-0.102	-0.074	0.102	2
OLSSkew	-0.132	-0.140	-0.134	-0.140	-0.102	0.030	5
LogitSkew	-0.173	-0.140	-0.132	-0.116	-0.079	0.093	3
HistoricalSkew	-0.129	-0.132	-0.132	-0.129	-0.124	0.004	6
MaxSkew	-0.164	-0.148	-0.130	-0.111	-0.093	0.071	4
Panel C: Five-Year Ahead Realized Skewness							
$QRSkew_{t+1,t+60}$	-0.076	-0.064	-0.060	-0.050	-0.038	0.038	1
$NaiveQRSkew_{t+1,t+60}$	-0.076	-0.063	-0.057	-0.051	-0.040	0.036	3
OLSSkew	-0.059	-0.061	-0.064	-0.066	-0.051	0.009	5
LogitSkew	-0.076	-0.066	-0.061	-0.055	-0.038	0.038	2
HistoricalSkew	-0.060	-0.062	-0.062	-0.059	-0.056	0.004	6
MaxSkew	-0.072	-0.070	-0.062	-0.053	-0.043	0.028	4

Table 7
Cross-Sectional NLS Regressions of Realized Skewness on Various Skewness Forecasts
This table shows the results from the following cross-sectional NLS regression:

$$RealizedSkew = \frac{1}{1+\sum_i e^{\beta_i}} QRSkew + \frac{e^{\beta_1}}{1+\sum_i e^{\beta_i}} SkewnessForecast_1 \left(+ \frac{e^{\beta_2}}{1+\sum_i e^{\beta_i}} SkewnessForecast_2 + \ldots \right) + \epsilon,$$

where $RealizedSkew$ measures realized stock return skewness over the next month ($RealizedSkew_{t+1}$; Panel A), the next twelve months ($RealizedSkew_{t+1,t+12}$; Panel B), or the next five years ($RealizedSkew_{t+1,t+60}$; Panel C); $QRSkew$ is a quantile regression based skewness forecast calculated from the same stock return interval as the $RealizedSkew$ variable with which it is associated; and $SkewnessForecast \in (OLSSkew, LogitSkew, HistoricalSkew, MaxSkew)$. $OLSSkew$ is an OLS-based forecast of the realized skewness of daily stock returns over the next 60 months; $LogitSkew$ is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; $HistoricalSkew$ is the realized daily stock return skewness over the previous 60 months; and $MaxSkew$ is the maximum daily return over the previous month. See Table A.1 for details. We convert all endogenous and exogenous variables into standard normal variables before running a cross-sectional regression. As exogenous variables, we use $QRSkew$ and one other skewness forecast or $QRSkew$ and all others. β_1, β_2, ... are free parameters, and ϵ is the residual. We estimate the models separately by month. Panels A, B, and C show the mean, the median, and the interquartile range of the variable coefficients, respectively. The sample period is December 1987 to December 2010.

	$QRSkew$=$QRSkew_{t+1}$					$QRSkew$=$QRSkew_{t+1,t+12}$					$QRSkew$=$QRSkew_{t+1,t+60}$				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Panel A: Mean Estimate															
QRSSkew	0.54	0.54	0.55	0.52	0.27	0.58	0.57	0.59	0.59	0.28	0.58	0.55	0.60	0.56	0.27
OLSSkew	0.46				0.20	0.42				0.24	0.42				0.23
LogitSkew		0.46			0.21		0.43			0.22		0.45			0.22
HistoricalSkew			0.45		0.08			0.41		0.10			0.40		0.10
MaxSkew				0.48	0.25				0.41	0.18				0.44	0.17
Panel B: Median Estimate															
QRSSkew	0.54	0.54	0.55	0.52	0.27	0.58	0.56	0.60	0.57	0.27	0.59	0.55	0.60	0.56	0.28
OLSSkew	0.46				0.20	0.42				0.23	0.41				0.24
LogitSkew		0.46			0.20		0.44			0.20		0.45			0.22
HistoricalSkew			0.45		0.07			0.40		0.08			0.40		0.10
MaxSkew				0.48	0.25				0.43	0.21				0.44	0.18
Panel C: Interquartile Range															
QRSSkew	0.08	0.09	0.06	0.08	0.08	0.08	0.13	0.08	0.11	0.08	0.06	0.08	0.04	0.08	0.10
OLSSkew	0.08				0.08	0.08				0.08	0.06				0.09
LogitSkew		0.09			0.11		0.13			0.19		0.08			0.10
HistoricalSkew			0.06		0.09			0.08		0.16			0.04		0.08
MaxSkew				0.08	0.10				0.11	0.09				0.08	0.09

Table 8
Portfolio Formation Exercises Based on Skewness Forecasts
The table reports the mean returns of portfolios sorted according to the various skewness forecasts. Panel A uses quantile-regression based skewness forecasts implied from fitted moments as sorting variables; Panel B uses those directly calculated from quantile estimates. In both panels, the subscripts indicate the length of the period over which the stock return underlying the skewness forecast is based. Panel C uses the alternative skewness forecasts as sorting variable. *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. See Table A.1 for more details. We form the portfolios as follows: At the end of every month t in our sample period, we sort stocks into portfolios according to the quintile breakpoints of one of the skewness forecasts. Portfolio 1 (Low) contains stocks with low skewness forecast values; portfolio 5 contains stocks with high skewness forecast values. We value-weight the portfolios and hold them over the next month. We also form a spread portfolio long on portfolio 5 and short on portfolio 1 ("5–1"). We report the mean return, the Q-factor model alpha, and the Fama-French (2015) five-factor model alpha of the spread portfolio under "raw," "Q," and "FF5," respectively. The alphas are computed from time-series regressions of the spread portfolio return on the relevant benchmark factors. Mean returns are per month and in percentage; t-statistics are calculated using the Newey and West (1987) formula with a lag length of twelve months and are in square parentheses. The sample period is December 1987 to December 2010.

	Skewness Forecast Deciles					5–1		
Skewness Forecast	1 (Low)	2	3	4	5 (High)	raw	Q	FF5
Panel A: Quantile Regression-Based Skewness Forecasts								
$QRSkew_{t+1}$	0.930	0.948	0.910	0.859	0.681	-0.250	0.193	-0.042
	[3.38]	[2.28]	[2.03]	[1.59]	[1.20]	[-0.59]	[1.00]	[-0.21]
$QRSkew_{t+1,t+12}$	0.937	1.002	1.201	0.918	0.488	-0.449	0.169	-0.046
	[3.62]	[2.55]	[2.46]	[1.59]	[0.74]	[-0.84]	[0.48]	[-0.20]
$QRSkew_{t+1,t+60}$	0.888	1.113	1.144	1.177	1.103	0.215	0.563	0.498
	[3.53]	[3.20]	[2.59]	[2.55]	[1.78]	[0.44]	[1.32]	[1.14]
Panel B: Nave Quantile Regression-Based Skewness Forecasts								
$NaiveQRSkew_{t+1}$	0.990	0.961	1.101	0.923	0.844	-0.147	0.057	0.012
	[2.70]	[2.89]	[3.29]	[2.90]	[3.26]	[-0.61]	[0.26]	[0.06]
$NaiveQRSkew_{t+1,t+12}$	0.946	1.009	1.287	0.927	0.340	-0.606	-0.221	-0.237
	[3.58]	[2.60]	[2.69]	[1.55]	[0.52]	[-1.22]	[-0.54]	[-0.88]
$NaiveQRSkew_{t+1,t+60}$	0.911	0.984	1.031	0.958	0.981	0.071	0.324	0.276
	[3.25]	[2.65]	[2.55]	[2.32]	[2.02]	[0.21]	[1.22]	[1.21]
Panel C: Other Skewness Forecasts								
OLSSkew	1.094	0.795	0.840	0.677	0.390	-0.704	-0.279	-0.562
	[3.76]	[2.26]	[2.26]	[1.75]	[0.82]	[-2.02]	[-0.97]	[-1.65]
LogitSkew	0.931	1.078	0.863	0.762	0.328	-0.602	-0.117	-0.277
	[3.28]	[2.80]	[1.75]	[1.37]	[0.51]	[-1.17]	[-0.49]	[-0.97]
HistoricalSkew	1.057	0.958	0.920	0.782	0.766	-0.290	-0.205	-0.209
	[3.86]	[3.10]	[3.01]	[2.14]	[2.13]	[-1.82]	[-1.51]	[-1.52]
MaxSkew	1.003	0.968	1.011	0.767	0.443	-0.560	-0.045	-0.189
	[4.39]	[3.31]	[2.74]	[1.42]	[0.72]	[-1.11]	[-0.13]	[-0.91]

Table 9
Spreads From Portfolio Formation Exercises Based on Skewness Forecast Combinations
The table reports the mean returns of spread portfolios long on stocks with high values for various skewness-forecast combinations and short on stocks with low values for them. The column labels indicate the quantile-regression based skewness forecast used in the combination; the row labels indicate with which other skewness forecast(s) the quantile-regression based forecast is combined. To combine the skewness forecasts, we estimate the cross-sectional NLS regression in Equation (20). As endogenous variable we use $RealizedSkew$ measured using the same return interval as $QRSkew$ and calculated with data running up to month t. We convert all regression variables into standard normal variables before estimation. We use the weights from the regression in conjunction with the month t values of the (standardized) skewness forecasts to calculate the forecast combination. The only exception occurs in row "all (equal-weights)," in which we report the results from equally-weighting the forecasts. *QRSkew* and *NaiveQRSkew* are quantile-regression based skewness forecasts; their subscripts indicate the length of the period over which the stock return underlying the measure is based. *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. See Table A.1 for details. We form the spread portfolios as follows: At the end of every month t in our sample period, we sort stocks into portfolios according to the quintile breakpoints of one of the forecast combinations. Portfolio 1 (Low) contains stocks with low forecast combination values; portfolio 5 contains stocks with high forecast combination values. We value-weight the portfolios and hold them over the next month. The spread portfolio is long on portfolio 5 and short on portfolio 1. We report the mean return, the Q-factor model alpha, and the Fama-French (2015) five-factor model alpha of the spread portfolio under "Raw Spread," "Q-Adjusted Spread," and "FF5-Adjusted Spread," respectively. The alphas are computed from time-series regressions of the spread portfolio return on the relevant benchmark factors. Mean returns are per month and in percentage; t-statistics (in square parentheses) are calculated using the Newey and West (1987) formula with a lag length of twelve months. The sample period is December 1987 to December 2010.

	Raw Spread			Q-Adjusted Spread			FF5-Adjusted Spread		
Combined With:	$QRSkew_{t+1}$	$QRSkew_{t+1,t+12}$	$QRSkew_{t+1,t+60}$	$QRSkew_{t+1}$	$QRSkew_{t+1,t+12}$	$QRSkew_{t+1,t+60}$	$QRSkew_{t+1}$	$QRSkew_{t+1,t+12}$	$QRSkew_{t+1,t+60}$
OLSSkew	-0.184	-0.389	-0.420	0.220	0.270	0.042	-0.045	-0.014	-0.038
	[-0.41]	[-0.78]	[-0.79]	[0.90]	[1.12]	[0.09]	[-0.14]	[-0.05]	[-0.13]
LogitSkew	-0.236	-0.456	-0.482	0.451	0.305	0.180	0.074	-0.002	0.043
	[-0.46]	[-0.82]	[-0.79]	[1.71]	[1.11]	[0.34]	[0.26]	[-0.01]	[0.13]
HistoricalSkew	-0.057	-0.282	0.134	0.195	0.167	0.694	0.033	0.053	0.626
	[-0.21]	[-0.69]	[0.27]	[1.51]	[0.68]	[1.74]	[0.31]	[0.42]	[2.97]
MaxSkew	-0.241	-0.499	0.225	0.253	0.172	0.588	-0.022	-0.020	0.538
	[-0.50]	[-0.89]	[0.46]	[1.17]	[0.48]	[1.39]	[-0.09]	[-0.09]	[2.28]
all (reg-weights)	-0.290	-0.509	-0.665	0.383	0.123	0.081	0.056	-0.108	-0.076
	[-0.59]	[-0.97]	[-1.09]	[1.54]	[0.49]	[0.17]	[0.21]	[-0.48]	[-0.32]
all (equal-weights)	-0.367	-0.483	-0.625	0.106	0.065	0.008	-0.116	-0.139	-0.076
	[-0.83]	[-1.00]	[-1.21]	[0.54]	[0.27]	[0.03]	[-0.64]	[-0.71]	[-0.32]

Table 10

Fama-MacBeth Regressions of Future Returns on Skewness Forecasts and Controls

The table shows the results from Fama-MacBeth (1973) regressions of stock returns on different skewness forecasts (*SkewForecast*) and control variables. *SkewForecast* is either *QRSkew*, *NaiveQRSkew*, *OLSSkew*, *LogitSkew*, *HistoricalSkew*, or *MaxSkew*. *QR-Skew* and *NaiveQRSkew* are quantile-regression based skewness forecasts; their subscripts indicate the length of the period over which the stock return underlying the measure is based. *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. The control variables are *MarketBeta*, *BookToMarket*, *Momentum*, *AssetGrowth*, *Profitability*, and *Volatility*. See Table A.1 for details. Estimates (in bold) are per month and in percentage. T-statistics (in square parentheses) are calculated using the Newey and West (1987) formula with a lag length of twelve months and are in square parentheses. The sample period is December 1987 to December 2010.

	SkewForecast =									
	QRSkew $_{t+1}$	*QRSkew* $_{t+1,t+12}$	*QRSkew* $_{t+1,t+60}$	*Naive QRSkew* $_{t+1}$	*Naive QRSkew* $_{t+1,t+12}$	*Naive QRSkew* $_{t+1,t+60}$	*OLS Skew*	*Logit Skew*	*Historical Skew*	*MAX Skew*
SkewForecast	**0.288**	**0.240**	**0.106**	**-2.380**	**2.559**	**3.252**	**-0.198**	**-0.248**	**-0.030**	**-4.879**
	[0.84]	[0.90]	[0.88]	[-1.32]	[1.21]	[1.37]	[-1.26]	[-1.22]	[-0.73]	[-5.77]
MarketBeta	**0.061**	**0.038**	**0.055**	**0.067**	**0.042**	**0.055**	**0.069**	**0.074**	**0.060**	**0.075**
	[0.76]	[0.51]	[0.72]	[0.82]	[1.21]	[0.66]	[0.78]	[1.00]	[0.71]	[0.90]
Size	**-0.065**	**-0.043**	**-0.076**	**-0.091**	**-0.053**	**-0.075**	**-0.131**	**-0.195**	**-0.108**	**-0.119**
	[-1.37]	[-1.48]	[-1.86]	[-2.64]	[-1.12]	[-1.78]	[-3.00]	[-2.47]	[-2.98]	[-3.27]
BookToMarket	**0.253**	**0.304**	**0.291**	**0.287**	**0.290**	**0.261**	**0.240**	**0.222**	**0.252**	**0.246**
	[2.86]	[3.92]	[3.92]	[2.99]	[3.63]	[3.04]	[2.62]	[2.59]	[2.73]	[2.68]
Momentum	**0.002**	**0.002**	**0.002**	**0.002**	**0.002**	**0.002**	**0.002**	**0.003**	**0.002**	**0.002**
	[0.73]	[0.63]	[0.63]	[0.71]	[0.70]	[0.66]	[0.54]	[1.19]	[0.62]	[0.58]
AssetGrowth	**-0.459**	**-0.472**	**-0.522**	**-0.475**	**-0.541**	**-0.489**	**-0.506**	**-0.559**	**-0.516**	**-0.513**
	[-3.42]	[-3.43]	[-3.57]	[-3.27]	[-3.54]	[-3.79]	[-3.45]	[-3.41]	[-3.49]	[-3.46]
Profitability	**0.103**	**0.104**	**0.088**	**0.081**	**0.085**	**0.083**	**0.082**	**0.138**	**0.087**	**0.085**
	[1.65]	[1.68]	[1.26]	[1.08]	[1.15]	[1.14]	[1.17]	[2.10]	[1.21]	[1.16]
Volatility	**-2.299**	**-2.443**	**-2.210**	**-2.307**	**-2.336**	**-2.317**	**-2.146**	**-1.302**	**-2.286**	**-1.137**
	[-3.84]	[-4.34]	[-3.76]	[-3.44]	[-3.95]	[-3.69]	[-3.30]	[-2.48]	[-3.47]	[-1.73]
Constant	**2.121**	**1.317**	**2.176**	**2.663**	**1.834**	**2.222**	**3.241**	**2.447**	**2.783**	**3.084**
	[2.61]	[1.52]	[2.79]	[4.09]	[2.20]	[3.06]	[4.07]	[3.44]	[4.25]	[4.68]
Adj. R^2	0.041	0.043	0.043	0.040	0.043	0.041	0.040	0.040	0.039	0.040

Table 11
Fama-MacBeth Regressions of Future Returns on Skew Forecast Combinations and Controls
The table shows the risk premium estimates of various skewness forecast combinations. We obtain the estimates from Fama-MacBeth (1973) regressions of stock returns on the skewness forecast combinations and control variables. The column labels indicate the quantile-regression based skewness forecast used in the combination; the row labels indicate with which other skewness forecast(s) the quantile-regression based forecast is combined. To combine the skewness forecasts, we estimate the cross-sectional NLS regression in Equation (20). As endogenous variable we use $RealizedSkew$ measured using the same return interval as $QRSkew$ and calculated with data running up to month t. We convert all regression variables into standard normal variables before estimation. We use the weights from the regression in conjunction with the month t values of the (standardized) skewness forecasts to calculate the forecast combination. The only exception occurs in row "all (equal-weights)," in which we report the results from equally-weighting the forecasts. *QRSkew* is a quantile-regression based skewness forecast; its subscript indicates the length of the period over which the stock return underlying the measure is based. *OLSSkew* is an OLS-based forecast of the realized daily stock return skewness over the next 60 months; *LogitSkew* is the fitted value from a logit model explaining a dummy variable equal to one if a stock's return over the next twelve months exceeds 100% and else zero; *HistoricalSkew* is the realized daily stock return skewness over the previous 60 months; and *MaxSkew* is the maximum daily return over the previous month. The control variables are $MarketBeta$, $BookToMarket$, $Momentum$, $AssetGrowth$, $Profitability$, and $Volatility$. See Table A.1 for details. Estimates (in bold) are per month and in percentage. T-statistics (in square parentheses) are calculated using the Newey and West (1987) formula with a lag length of twelve months. The sample period is December 1987 to December 2010.

Combined With:	$QRSkew_{t+1}$	$QRSkew_{t+1,t+12}$	$QRSkew_{t+1,t+60}$
OLSSkew	**-0.029**	**0.366**	**0.218**
	[-0.10]	[1.21]	[0.91]
LogitSkew	**0.077**	**0.272**	**-0.015**
	[0.24]	[0.78]	[-0.08]
HistoricalSkew	**-0.016**	**0.153**	**0.089**
	[-0.13]	[0.92]	[0.64]
MaxSkew	**-0.172**	**0.384**	**0.129**
	[-0.26]	[1.43]	[0.61]
all others (reg-weights)	**0.422**	**0.574**	**0.337**
	[1.02]	[1.46]	[1.01]
all others (equal-weights)	**0.146**	**0.294**	**0.249**
	[0.53]	[1.01]	[0.83]
Controls	YES	YES	YES