Journal of Financial Economics 135 (2020) 399–424

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Measuring skewness premia☆

Hugues Langlois

HEC Paris, 1 rue de la Libération, Jouy-en-Josas 78350, France

ARTICLE INFO

Article history:
Received 18 June 2018
Revised 14 November 2018
Accepted 13 December 2018
Available online 14 June 2019

JEL classification:
G12

Keywords:
Systematic skewness
Coskewness
Idiosyncratic skewness
Large panel regression
Forecasting

ABSTRACT

We provide a new methodology to empirically investigate the respective roles of systematic and idiosyncratic skewness in explaining expected stock returns. Using a large number of predictors, we forecast the cross-sectional ranks of systematic and idiosyncratic skewness, which are easier to predict than their actual values. Compared to other measures of ex ante systematic skewness, our forecasts create a significant spread in ex post systematic skewness. A predicted systematic skewness risk factor carries a significant and robust risk premium that ranges from 6% to 12% per year. In contrast, the role of idiosyncratic skewness in pricing stocks is less robust.

1. Introduction

Stocks with positively skewed returns are attractive to most investors because they occasionally pay large returns. Stocks with negatively skewed returns are less attractive because they sometimes drastically fall in value. However, not all kinds of return skewness are equal. Stocks with higher systematic return skewness are appealing because they offer defensive returns during bad times; these stocks provide downside protection. On the other hand, stocks with positive idiosyncratic skewness are sought for their potential for high returns regardless of broad market movements; these stocks provide a lottery payoff.

In studying the theoretical link between skewness and asset prices, previous research has proposed models in which only systematic skewness carries a risk premium (see Rubinstein, 1973; Kraus and Litzenberger, 1976; Harvey and Siddique, 2000) and models in which total skewness—both systematic and idiosyncratic—are important in pricing securities (see Brunnermeier et al., 2007; Barberis and Huang, 2008; Mitton and Vorkink, 2007).[1] These models mainly differ by their assumptions about investors' preferences.

In this paper, we empirically investigate the respective importance of systematic and idiosyncratic skewness in explaining differences in expected returns across stocks. Answering this question requires measuring each type of skewness. While measuring return skewness is a challenging task, distinguishing between types of skewnesses is harder still. Two broad methodologies exist. First, one can use past return skewness to predict its future value. But

☆ I am thankful to Daniel Andrei, Bruno Biais, Martin Boyer, Tolga Cenesizoglu, Jean-Edouard Colliard, Robert Dittmar (the referee), Christian Dorion, Thierry Foucault, Mathieu Fournier, René Garcia, Elise Gourier, Denis Gromb, Johan Hombert, Alexandre Jeanneret, Jiro Kondo, Augustin Landier, Hayne Leland, Ricardo Lopez Aliouchkin (discussant), Jacques Olivier, Ioanid Rosu, Guillaume Roussellet, Bill Schwert (the editor), Christophe Spaenjers, Anders Trolle, seminar participants at HEC Paris, McGill University, HEC Montreal, and University College Dublin for helpful comments and discussions. This research is supported by a grant of the French National Research Agency (ANR), Investissements d'avenir (LabEx Ecodec/ANR-11-LABX-0047).

E-mail address: langlois@hec.fr

[1] For systematic skewness, see also Simaan (1993), Dittmar (2002), Dahlquist et al. (2017), and Chabi-Yo et al. (2014).

https://doi.org/10.1016/j.jfineco.2019.06.002
0304-405X/

given the low time-series persistence of skewness measures, this methodology is bound to generate poor results. In their textbook treatment of this literature, Bali et al. (2016, p.330) state that “these variables are likely to be very noisy, perhaps to the point of being ineffective at measuring the characteristics of the stock that they are designed to capture.”

A second methodology is to measure skewness from option prices, but this approach also suffers from limitations. The risk-neutral skewness obtained from options needs to be translated to physical skewness at the cost of some assumption on how the two are related. Similarly, extracting the systematic part of option-implied skewness requires further assumptions.[2] In addition, given the availability of option data, empirical tests are restricted to a shorter period and a smaller cross-section than when using only returns.[3]

This paper develops a novel methodology for predicting differences in systematic and idiosyncratic skewness across stocks. We use our methodology to build a predicted systematic skewness factor and idiosyncratic skewness sorted portfolios as well as to describe which variables best predict each type of skewness. The predicted systematic skewness factor captures future systematic skewness risk better than other measures, is distinct from leading equity risk factors, and carries a significant risk premium. In contrast, we find weaker evidence that predicted idiosyncratic skewness is priced in US stocks.

Our results are important because previous research has found that idiosyncratic skewness has a more robust pricing importance than systematic skewness. When comparing systematic to idiosyncratic skewness, different empirical methodologies are often used to measure each type of risk. Instead, we use the same empirical methodology to predict each measure, show that we successfully forecast their respective future realized values, and find that systematic skewness, in contrast to idiosyncratic skewness, has a robust and economically large premium.

We make three main contributions. First, we develop a novel methodology to forecast differences in systematic skewness between stocks. Buying a diversified portfolio of stocks with low measures of systematic skewness and short-selling a diversified portfolio of stocks with high values should isolate systematic skewness risk. Unfortunately, sorting stocks on past systematic skewness measures to create a low-minus-high factor produces insignificant realized, i.e. ex post, systematic skewness. These realized systematic skewness measures are even significantly positive in some cases, the opposite sign of what these factors are designed to produce.[4]

Our main insight is that forming a low-minus-high systematic skewness factor only requires the cross-sectional ordering of stock systematic skewness, not their actual values. Therefore, we predict the former, which is considerably easier than predicting the latter. We use large panel regressions to predict the future cross-sectional ranks of individual stock systematic skewness using the cross-sectional ranks of a large number of past risk measures and firm characteristics. We form each month a portfolio that buys stocks with low predicted systematic skewness cross-sectional ranks and short-sells stocks with high predicted systematic skewness cross-sectional ranks. We find that this predicted systematic skewness (*PSS*) factor generates, from July 1963 to December 2017, a significantly positive average excess return of 5.37% per year with a Sharpe ratio of 0.38. The corresponding values for the market portfolio are 6.37% and 0.42, respectively. Most importantly, the *PSS* factor has a significantly negative realized systematic skewness.

Despite the large number of predictors used, we obtain stable regression coefficients over time that increase our confidence in our results. This result is surprising given that using multiple predictors in a predictive model—the kitchen sink approach (see Goyal and Welch, 2008)—generally performs badly when predicting expected returns. The key feature of our approach is that we use cross-sectional ranks of predictors and the dependent variable. For example, we do not ask whether a large market capitalization predicts a high systematic skewness. Rather, we estimate whether being among the largest firms is related to having the highest systematic skewness across stocks. More formally, we show analytically and by simulation that the better forecasting performance of our methodology comes from removing the impact of the time-varying marginal distributions of the predictors and predicted variable on the estimated regression coefficients.

Our second contribution is to use the *PSS* factor in formal asset pricing tests. In the main exercise, we use 25 size and book-to-market ratio sorted portfolios, 25 size and momentum sorted portfolios, and 25 size and predicted systematic skewness sorted portfolios of US stocks as test assets. We test different models: the capital asset pricing model (CAPM) with the market factor; the Fama–French–Carhart four-factor model with market, size, value, and momentum factors (Fama and French, 1993; Carhart, 1997); and the five-factor model with market, size, value, profitability, and investment factors (Fama and French, 2015). We also test another model designed to capture downside risk, the downside-beta CAPM of Ang et al. (2006) and Lettau et al. (2014). To assess the value of adding the *PSS* factor to each model, we run time-series regressions of test portfolio excess returns on the factors and run a cross-sectional regression of their average returns on their factor exposures to estimate risk premia. Remarkably,

[2] Following Bakshi et al. (2003), Conrad et al. (2013) assume a one-factor structure under the risk-neutral measure to recover risk-neutral systematic skewness from option-implied moments of individual stocks and of the market portfolio. Because there are no options traded on the market portfolio implied by theoretical models, they use S&P 500 Index options as a proxy. Schneider et al. (2017) rely on a structural model of levered firms to empirically relate option-implied skewness to systematic skewness.

[3] Data for US equity option on Optionmetrics start in 1996. Conrad et al. (2013) report an average of 92 stocks in their bottom and top portfolios that contain stocks with the lowest 30% and the highest 30% option-implied skewness, respectively. Schneider et al. (2017) use an average of 1,800 US stocks from January 1996 to August 2014, whereas there are, on average, 5,361 stocks during the same period.

[4] We explore three different systematic skewness measures that we estimate using either monthly or daily returns.

in all but one model the *PSS* risk premium is positive and significant, ranging from 0.47% to 0.98% per month.

The *t*-ratios of the *PSS* factor range from 2.33 to 3.06, which is impressive for several reasons. First, given the growing number of risk factors that have been tested over the years, many of them not based on an economic model, Harvey et al. (2016) advocate using a higher standard than the traditional *t*-ratio of two when assessing the value of a new risk factor. In this paper, however, we do not propose a new risk factor but rather a new methodology to measure a traditional risk factor that has been tested as far back as in Kraus and Litzenberger (1976). Its risk premium is supported by an economic model in which investors require a compensation for bearing lower systematic skewness.[5] Second, we obtain *t*-ratios using the methodology of Kan et al. (2013) that accounts for model misspecification. Finally, we find that the *PSS* risk premium is significant even when factors designed to capture the cross-sectional variation in average returns of the test assets are included in our tests (e.g., size and value factors with size and value sorted test portfolios).

As a robustness check, we also estimate the *PSS* risk premium using the methodology of Giglio and Xiu (2017), which accounts for omitted risk factors and measurement error in the risk factor of interest. Using either the size, book-to-market ratio, momentum, and skewness portfolios, or the 202 test portfolios used in Giglio and Xiu (2017), we find that the risk premium for *PSS* is always significantly positive.

Our third contribution is to examine the extent to which idiosyncratic skewness is priced in US stocks. We use our panel regression model to predict the cross-sectional ranks of future idiosyncratic skewness. Stocks with low predicted idiosyncratic skewness ranks have lower realized idiosyncratic skewness than stocks with high predicted ranks, and the spread in realized idiosyncratic skewness is larger than when using other idiosyncratic skewness predictors.

Then, we form equal- and value-weighted portfolios of stocks sorted by their predicted idiosyncratic skewness rank, and a long-short portfolio that buys low idiosyncratic skewness stocks and short-sells high idiosyncratic skewness stocks. We run time-series regressions of their returns on different factor models that include the *PSS* factor. Our first result is that *PSS* has a strong explanatory power for idiosyncratic skewness sorted portfolios: low idiosyncratic skewness stocks are negatively exposed to the *PSS* factor. However, the risk-adjusted performance of the low-minus-high idiosyncratic skewness portfolio is significant in some of the specifications, in particular for value-weighted portfolios. Hence, the *PSS*, while important, is not sufficient to fully explain the idiosyncratic skewness effect. When controlling for other risk factors such as momentum and profitability, the idiosyncratic skewness risk-adjusted performances are never significant. Our results are robust to predicting total instead of idiosyncratic skewness and to using an alternative quantile-based measure of either idiosyncratic or total skewness.

Our methodology identifies what are the determinants of systematic and idiosyncratic skewness. Except for higher momentum, higher price impact, and lower beta that predict both lower systematic and idiosyncratic skewness, predictors of systematic skewness are different from those of idiosyncratic skewness. For example, we confirm previous findings that skewness is negatively related to firm size.[6] Hence, it is a poor candidate to explain the size effect: small firms have higher average returns than large firms but also higher skewness that should be accompanied by lower average returns. However, large firms also have higher systematic skewness compared to small firms.[7] Consequently, we find that once we control for the *PSS* factor, the size factor is no longer needed in asset pricing tests.

To ensure that systematic skewness is not just a size effect in disguise, we provide two additional analyses. First, we use double-sorted portfolios to show that there are significant differences in average return across systematic skewness sorted portfolios within each size quintile. Conversely, the differences in average returns between the top and the bottom size-sorted quintile portfolios within each systematic skewness quintile are almost never significant. Second, we show that the significant α obtained by regressing the size factor on other factors in the Fama–French five-factor model disappears when *PSS* is added to the model. In contrast, the *PSS* α remains significant regardless of whether the size factor is included in the five-factor model. Hence, despite their positive correlation, *PSS* contains pricing information that is distinct from the size factor.

Another example of the difference in skewness predictors is idiosyncratic volatility. A higher cross-sectional rank of idiosyncratic volatility predicts a lower systematic skewness rank but also a higher idiosyncratic skewness rank. Boyer et al. (2010) (hereafter BMV) show that past idiosyncratic volatility is a stronger predictor of idiosyncratic skewness than past idiosyncratic skewness. In contrast, we find that the lagged cross-sectional rank of idiosyncratic skewness is a stronger predictor of its future value than lagged idiosyncratic volatility rank, with an average coefficient twice as large.

Our work is related to different strands of literature. There is a large literature on the link between return skewness and asset prices; see Bali et al. (2016) (Ch. 14) for a review. Models with well-diversified expected utility maximizing investors imply a compensation for negative systematic coskewness, which measures the contribution of an asset to the skewness of the market portfolio (see Rubinstein, 1973; Kraus and Litzenberger, 1976; Harvey and Siddique, 2000; Chabi-Yo et al., 2014). See also Dittmar (2002) and Smith (2007) for additional empirical

[5] See Rubinstein (1973); Kraus and Litzenberger (1976); Harvey and Siddique (2000); and Chabi-Yo et al. (2014).

[6] See, among others, Chen et al. (2001); Boyer et al. (2010); and Conrad et al. (2013).

[7] Bali et al. (2016) also find that coskewness (skewness) is positively (negatively) correlated with firm size. See also Albuquerque (2012) for a model to explain the difference between stock- and market-level skewness.

evidence. We provide a new methodology to measure the coskewness risk premium.

Models with other types of investor preferences, in contrast, imply that total skewness of an individual stock is priced (see Brunnermeier et al., 2007; Barberis and Huang, 2008). Mitton and Vorkink (2007) also study preference for skewness in retail investor portfolios. Empirical evidence of the negative relation between total stock skewness and returns implied by these models are provided in BMV and using a variable related to idiosyncratic skewness in Bali et al. (2011). Our results provide empirical evidence to compare both types of models using return skewnesses estimated using the same methodology.

Our work is also related to papers that use past return skewness measures and firm characteristics to predict skewness. Chen et al. (2001) (CHS hereafter) and BMV run cross-sectional regressions to predict future total and idiosyncratic skewness, respectively, using past risk measures and firm characteristics.[8] We run cross-sectional regressions to predict the ordering of individual stock skewness, not their values. We also use panel regressions to predict the systematic part of skewness. Zhang (2005) groups stocks with similar characteristics and uses all their recent returns to compute a stock's skewness. In our panel regression framework, stocks with similar characteristics will have similar forecasted skewness ranks. Freyberger et al. (2017) use the adaptive group LASSO to determine which stock characteristic ranks have explanatory power for stocks returns. We use stock characteristic ranks to predict skewness measures instead of returns and then build a risk factor suggested by an asset pricing model. Using intra-day returns from 1993 to 2013, Amaya et al. (2015) show that realized skewness negatively predicts future returns. We relate future skewness to both past risk measures and firm characteristics and also use our methodology to predict systematic skewness.

Our work is also related to papers using option-implied information about third moments. Christoffersen et al. (2017) show that the conditional price of coskewness risk can be obtained from the index option-implied variance risk premium. Our approach uses standard cross-sectional regressions to estimate the risk premia of several risk factors and shows that the price of coskewness risk is robustly significant. Conrad et al. (2013) show that future stock returns are negatively related to their option-implied skewness but not to risk-neutral coskewness. Whereas they average daily option-implied skewness within a quarter, Rehman and Vilkov (2012) find a positive relation with future returns using the latest daily option-implied skewness. Schneider et al. (2017) use option-implied skewness to explain returns on low-risk anomalies.[9] Our methodology does not require option data and uses a longer sample period and larger cross-section of stocks (3261 stocks, on average, from 1963 to 2017) to disentangle the role of systematic and idiosyncratic skewness. Whereas we focus on the cross-section of stock expected returns, Bali and Murray (2013) and Boyer and Vorkink (2014) find that stock-option portfolio and option returns, respectively, are negatively related to option-implied skewness.[10]

More broadly, our paper is related to the literature on the importance of return skewness for portfolio choice. Dahlquist et al. (2017) theoretically derive the optimal portfolio choice of a generalized disappointment aversion investor in the presence of return skewness and show that it can explain patterns in empirically observed asset allocation choices. Ghysels et al. (2016) empirically show that skewness is an important factor in allocating a portfolio across emerging market equity indexes. DeMiguel et al. (2013) use option-implied skewness as a predictor of stock returns to improve the performance of mean–variance portfolios. Our results provide an estimate of the cost of hedging coskewness risk, or alternatively, the extra return an investor earns by bearing coskewness risk.

Section 2 presents a new systematic skewness factor, Section 3 runs asset pricing tests, Section 4 analyzes the performance of idiosyncratic skewness portfolios, Section 5 provides Monte Carlo simulations to show the higher predictive ability of our methodology, and Section 6 concludes.

2. Measuring systematic skewness

In this section, we start by discussing the asset pricing implication of systematic skewness. We then show in Section 2.2 the inability of past measures of systematic skewness to capture future skewness risk that will motivate our methodology introduced in Section 2.3. Finally, we discuss the performance of our new systematic skewness factor in Section 2.4, explain why it outperforms in Section 2.5, and describe the predictors of systematic skewness in Section 2.6.

2.1. How does systematic skewness affect risk premia?

We consider the conditional three-moment CAPM in which the expected return of stock i in month t, $r_{i,t}$, is given by

$$E_{t-1}\left[r_{i,t}-r_{f,t}\right]=\gamma_{M,t}Cov_{t-1}(r_{i,t},r_{M,t})+\gamma_{M^2,t}Cov_{t-1}\left(r_{i,t},r^2_{M,t}\right), \tag{1}$$

where $r_{f,t}$ is the risk-free rate for period t; $r_{M,t}$ is the market portfolio return for period t; $\gamma_{M,t}$ and $\gamma_{M^2,t}$ are, respectively, the time t prices of covariance and coskewness risk; and $E_{t-1}[\cdot]$ and $Cov_{t-1}(\cdot)$ denote, respectively, the expected value and covariance conditional on information at time $t-1$.

In this model, investors have a preference for higher portfolio return skewness. The contribution of a stock to the market portfolio skewness is captured by the coskewness term $Cov_{t-1}(r_{i,t},r^2_{M,t})$. For example, adding a stock with a negative coskewness to the market portfolio makes the latter more negatively skewed. The impact on market portfolio skewness is smaller when adding a stock with a less negative coskewness. In contrast, adding a positive

[8] See Cosemans et al. (2015) for a similar panel regression approach to predict stock betas using firm characteristics.

[9] Leland (1999) also shows that skewed strategies are misevaluated by the CAPM.

[10] See also Cremers and Weinbaum (2010) and Xing et al. (2010).

coskewness stock increases the market portfolio's skewness.

Investors require higher expected returns for holding stocks that decrease the market portfolio's skewness, and hence the price of coskewness risk, $\gamma_{M^2,t}$, is negative. The economic implication of the three-moment CAPM is that coskewness is the right measure of systematic skewness. We therefore use the term coskewness in the following sections to refer to the systematic part of skewness.

There are different ways of obtaining the three-moment CAPM. Using a Taylor expansion for marginal utility, Rubinstein (1973) and Kraus and Litzenberger (1976) impose some restrictions on investor preferences to obtain Eq. (1) in an unconditional setting. Harvey and Siddique (2000) obtain this model in a representative agent framework using either a quadratic stochastic discount factor or a Taylor expansion of marginal utility. Dittmar (2002) further considers cokurtosis and human capital in addition to coskewness.

Another way of obtaining a model related to Eq. (1) is to impose an assumption on the distribution of return shocks. Simaan (1993) provides a model with spherical shocks and one common nonspherical shock to create systematic skewness. Dahlquist et al. (2017) use Gaussian shocks and one common exponential shock to create systematic skewness. Both papers derive the asset pricing implications of systematic skewness in stock returns, the first in an expected utility framework and the second with generalized disappointment aversion preferences. Dahlquist et al. (2017) show that using generalized disappointment aversion instead of expected utility will lead to a larger importance for return skewness.

In the next section, we turn to the empirical task of building a risk factor to capture coskewness risk. In Section 4, we empirically examine whether idiosyncratic skewness is also priced.

2.2. Are systematic skewness measures persistent?

To test the three-moment CAPM in Eq. (1), we follow Harvey and Siddique (2000) and build a hedge portfolio to capture the source of risk that generates systematic skewness in asset returns. We form a well-diversified portfolio of stocks with low coskewness and a well-diversified portfolio of stocks with high coskewness. The return spread between the former and the latter captures the coskewness factor shock.[11] We show in this section that coskewness estimated on past returns is a weak predictor of future coskewness and sorting stocks on past coskewness will fail to capture the systematic skewness risk premium.

We consider several measures of systematic skewness that have been used in past studies. First, we denote as $Cos_{i,t}$ the coskewness from Eq. (1)

$$Cos_{i,t} = Cov_{t-1}\left(r_{i,t}, r_{M,t}^2\right). \tag{2}$$

Second, we use the regression coefficient $\beta_{M^2,i,t}$ in a regression of excess stock returns on a constant, the excess market return, and the squared excess market return. This measure is motivated by the beta representation of the three-moment CAPM, $E[r_{i,t} - r_{f,t}] = \beta_{M,i,t}\mu_{M,t} + \beta_{M^2,i,t}\mu_{M^2,t}$, where $\mu_{M,t}$ and $\mu_{M^2,t}$ are the market and coskewness risk premia, respectively.

Third, we follow Harvey and Siddique (2000) and compute a measure of standardized coskewness, $\beta_{HS,i,t}$, for stock i defined as

$$\beta_{HS,i,t} = \frac{E_{t-1}\left[\epsilon_{i,t}\epsilon_{M,t}^2\right]}{\sqrt{E_{t-1}\left[\epsilon_{i,t}^2\right]E_{t-1}\left[\epsilon_{M,t}^2\right]}}, \tag{3}$$

where $\epsilon_{i,t} = r_{i,t} - r_{f,t} - \alpha_i - \beta_{M,i}(r_{M,t} - r_{f,t})$ is the residual from a regression of stock i's excess return on a constant and the market excess return, and $\epsilon_{M,t} = r_{M,t} - r_{f,t} - \mu_M$ is the deviation of market excess returns from its mean. The three measures $\text{Cos}_{i,t}$, $\beta_{M^2,i,t}$, and $\beta_{HS,i,t}$ are related through the specification of a regression of r on r_M^2 or on both r_M and r_M^2. The advantages of using $\beta_{HS,i,t}$ are that it is zero for the market portfolio, unit free, and akin to a factor loading. Because the market portfolio has a benchmark value of zero, we use $\beta_{HS,i,t}$ to measure and compare the realized systematic skewness of different risk factors constructed below. However, our conclusions are robust to using any other coskewness measure.

We use daily and monthly returns of all common stocks listed on NYSE, Amex, Nasdaq, and NYSE Arca from July 1963 to December 2017. Appendix A provides more details on our data construction. Each month, we compute the different coskewness measures for all available stocks. For each measure, we form two value-weighted portfolios: one containing the 30% of stocks with the lowest coskewness and one containing the top 30% stocks with the highest coskewness. The return on the factor is the return on the low-coskewness portfolio minus the return on the high-coskewness portfolio. The asset pricing relation in Eq. (1) implies that such a long-short portfolio should have positive average returns.

For each month t and each coskewness measure, we create two sets of risk factors. First, we use monthly returns over the last 60 months, from $t-60$ to $t-1$, to estimate each of the three systematic skewness measures. Second, we also investigate the added value of higher frequency data by using daily returns over the last year, from the first day of month $t-12$ to the last day of month $t-1$, to compute daily versions of the three measures. To measure coskewness, Harvey and Siddique (2000) use the monthly version of $\beta_{HS,i,t}$, and Bali et al. (2017) use the daily version of $\beta_{M^2,i,t}$.

We report their summary performance statistics in the first six rows of Table 1. We first report annualized average excess returns, volatilities, and Sharpe ratios. Across the coskewness factors, average excess returns range from −0.78% to 2.77%, but all are insignificant. The maximum Sharpe ratio across factors is 0.23, approximately half of the level for the market portfolio reported in the last line.

[11] Alternatively, we can use $r_{M,t}^2$ as a risk factor in asset pricing tests and estimate coskewness as the covariance between test portfolio returns and $r_{M,t}^2$. We instead predict coskewness at the stock level to form a low-minus-high coskewness factor. The main reason is that we want to use the same empirical methodology to predict systematic and idiosyncratic skewness to assess their relative asset pricing importance. Estimating idiosyncratic skewness using portfolios is not possible because portfolio aggregation diversifies idiosyncratic risk.

Table 1
Summary statistics for factors built from different coskewness measures.
We report summary statistics of monthly returns of different coskewness factors and the value-weighted market portfolio from July 1963 to December 2017. We report the annualized average return (in %), volatility (in %), and the Sharpe ratio. Next, we report the market β_M with the value-weighted market portfolio and the systematic skewness β_{HS} defined in Eq. (3). Finally, we report the regression α (annualized in %) using different factor models. Each month and for each coskewness measure, we compute the return of a factor that is long a value-weighted portfolio containing the stocks with the lowest 30% values and short a value-weighted portfolio containing the stocks with the top 30% values. Cos is the covariance between stock returns and squared market returns. β_{M^2} is the regression coefficient on squared market returns in a regression of excess stock returns on a constant, the market excess return, and its square. β_{HS} is the standardized coskewness measure of Harvey and Siddique (2000) defined in Eq. (3). We compute a monthly version of each of these three measures using monthly returns over the past 60 months and a daily version using daily returns over the past 12 months. Each month, we run the panel regression (4) to predict the cross-sectional rank of future coskewness using cross-sectional ranks of predictors on the right-hand side. We form the predicted systematic skewness factors *PSS* by forming a value-weighted portfolio with the stocks with the bottom 30% predicted coskewness and shorting a value-weighted portfolio with the stocks with the top 30% predicted coskewness. The *PSS* (unranked) uses a predictive panel regression without computing the cross-sectional ranks of future coskewness on the left-hand side. The market return is the value-weighted portfolio of all stocks on NYSE, Amex, and Nasdaq. For each factor and the market portfolio, we simulate 10,000 samples from a bivariate normal distribution with the same means and covariance matrix and compute the standardized coskewness to obtain the significance level. All other significance levels are obtained using a Newey–West estimator with $T^{0.25} \approx 6$ lags. * and ** denote significance at the 5% and 1% level, respectively.

Factor	Annualized average excess return	Annualized volatility	Sharpe ratio	β_M	Standardized systematic skewness β_{HS}	Annualized CAPM α	Annualized four-factor α	Annualized five-factor α
Monthly *Cos*	2.77	12.24	0.23	0.27**	0.32**	1.03	2.64	3.22*
Monthly β_{M^2}	1.80	9.45	0.19	0.04	0.12**	1.51	0.76	1.22
Monthly β_{HS}	1.89	8.55	0.22	0.00	0.14**	1.86	0.61	0.52
Daily *Cos*	1.46	12.13	0.12	0.16**	−0.01	0.41	−2.97	−0.76
Daily β_{M^2}	0.00	10.70	0.09	0.08	−0.02	0.52	−2.59	−1.36
Daily β_{HS}	−0.78	9.21	−0.08	−0.01	−0.04	−0.73	2.01	1.05
PSS	5.37**	14.19	0.38	0.21**	−0.27**	4.02*	−0.36	5.91**
PSS (unranked)	2.52	13.40	0.19	0.14	−0.16**	1.61	−0.95	4.55**
Market	6.37**	15.20	0.42		-			

Most importantly, we report in the sixth column for each factor their realized, i.e. ex post, standardized coskewness measure in Eq. (3). Factors built from monthly coskewness measures have significant realized coskewness but of the wrong sign (i.e., positive instead of negative). Therefore, past measures of coskewness do not create significantly negative spreads in ex post coskewness. Daily coskewness measures fare better; they have negative realized coskewness measures, but all are insignificant. These results are consistent with the low autocorrelations of coskewness measured at the stock level found in Bali et al. (2016) (see Table 14.7). In Section 6 of the Online Appendix, we show that we can create a significant spread in future covariance by sorting stocks on past daily covariance (as in Bali et al., 2017, for example). Hence, the lack of persistence problem is specific to coskewness.

The significant positive realized coskewness for monthly measures is surprising. In unreported results, we compute the contingency table of past stock coskewness computed over months $t-60$ to $t-1$ and the month t values of $r_{i,t} \times r_{M,t}^2$. We find that stocks in the bottom (top) 30% of past coskewness are equally likely to be in the bottom (top) 30% of $r_{i,t} \times r_{M,t}^2$ values as they are to move to the top (bottom) 30% group, which is indicative of the low time-series persistence of coskewness. Therefore, we should expect the coskewness of this factor to be close to zero. Instead, we find significantly positive coskewness values in Table 1 because of the impact of the October 1987 crash. During that month, the coskewness factors built using monthly measures of past coskewness experience positive returns, while the market factor plummets by more than 23%. Once we remove this monthly observation, the realized coskewness of these three factors are still positive but not significant anymore.

Another perspective is presented in the last three columns. We follow the insight of Barillas and Shanken (2016) and report the coskewness factors' α from multi-factor time-series regressions. Barillas and Shanken (2016) show that to judge the added value of a factor in explaining the cross-section of average returns, it is sufficient to show that it has a significant α when regressed against a benchmark set of risk factors.

We consider different models: the CAPM with the market factor (*MKT*); the Carhart–Fama–French four-factor model with *MKT*, size (*SMB*), value (*HML*), and momentum (*MOM*) factors; and the Fama–French five-factor model with *MKT*, the five-factor model's size factor SMB_{FF5}, *HML*, profitability (*RMW*), and investment (*CMA*) factors.[12] In all but one case, the α of coskewness factors are insignificant, suggesting that the coskewness factors do not add any explanatory power to existing factor models.

We can conclude from these results that past measures of coskewness are not persistent enough to generate a spread in future coskewness. In addition, the low-minus-high coskewness factors do not have positive risk-adjusted returns, either by themselves or when judged against other risk factors. Accordingly, we present in the next section a novel methodology to form the coskewness factor.

2.3. A predictive systematic skewness factor

In this section, we present a model for predicting the cross-sectional rank of conditional systematic skewness. Our methodology has three distinct features. First, we use a large number of risk measures and firm characteristics to

[12] All factor data are from Kenneth French's website.

predict future coskewness. Second, as the dependent variable, we use the cross-sectional ranks of stock coskewness, not their values. Third, we use the cross-sectional ranks of predictors. Given that the compositions of the long and short portfolios in a systematic skewness factor are determined by how stocks are ordered, we only need to forecast the cross-sectional ranks of stock coskewness values.

Each month t (say, January 2018), we run the following panel regression using all available stocks and historical data:

$$F\left(Cos_{i,k-12\to k-1}\right) = \kappa + F\left(Y_{i,k-24\to k-13}\right)\theta + F\left(X_{i,k-13}\right)\phi + \varepsilon_{i,k-12\to k-1}, \tag{4}$$

where $k = 25, 26, \ldots, t$; $i = 1, \ldots, N_k$; and N_k is the number of stocks available at time k. In this equation, $Cos_{i,k-12\to k-1}$ is the coskewness Cos of stock i from Eq. (2) computed using daily returns from month $k-12$ to month $k-1$ (e.g., January to December 2017); the K_Y variables $Y_{i,k-24\to k-13}$ are risk measures (volatility, β_M, coskewness, etc.) computed using stock i's daily returns from month $k-24$ to month $k-13$ (e.g., January to December 2016); the K_X-vector $X_{i,k-13}$ are characteristics (size, book-to-price ratio, momentum, etc.) of stock i observed at the end of month $k-13$; and $\varepsilon_{i,k-12\to k-1}$ are random shocks. We use a period of 12 months to measure risk measures such as coskewness because it provides a reasonable trade-off between having enough returns while allowing for variations over time.

The function $F(x_{i,t}) = \frac{Rank(x_{i,t})}{N_t+1}$ computes the normalized rank of variable $x_{i,t}$ in the cross-section of x_t. The $Rank(x_{i,t})$ function gives the order (1,2,...,N_t) of variable $x_{i,t}$ in all x_t values sorted in ascending order. We divide by $N_t + 1$ to obtain a variable that falls between 0 and 1.

The K_Y-by-one vector of coefficients θ and K_X-by-one vector of coefficients ϕ measure the ability of past ranks of risk measures and characteristics, respectively, to predict the future rank of a stock coskewness in the cross-section. κ is a constant. We run the panel regression (4) using all monthly observations from month 25 to month t. Each month, we use all stocks in the cross-section for which the values Cos, Y, and X are available. By estimating regression (4), we model how past cross-sectional ranks of risk measures and characteristics predict future coskewness ranks.[13] For example, we identify whether being among the largest market capitalization firms is associated with having an above median coskewness over the next 12 months.

To form our coskewness factor, we first compute the model predicted coskewness ranks for month t using our regression estimates, $\hat{\kappa}$, $\hat{\theta}$, and $\hat{\phi}$, as

$$F(\widehat{Cos_{i,t\to t+11}}) = \hat{\kappa} + F(Y_{i,t-12\to t-1})\hat{\theta} + F(X_{i,t-1})\hat{\phi}. \tag{5}$$

Finally, we form our coskewness factor as the return spread between the value-weighted portfolio containing the bottom 30% of stocks with the lowest predicted coskewness cross-sectional ranks and the value-weighted portfolio containing the top 30% of stocks with the highest predicted coskewness cross-sectional ranks. We denote the resulting factor as the predicted systematic skewness factor or PSS for short.[14]

In our empirical implementation, we use β_M, idiosyncratic volatility, coskewness Cos, and idiosyncratic skewness as past risk measures Y. These measures capture the systematic and idiosyncratic second and third order moments and therefore describe the shape of the distribution of past returns.

As firm characteristics, we use variables identified in the literature as being related to either average returns or to future skewness values. We use market capitalization and book-to-price ratio (Fama and French, 1993), profitability and investment (Fama and French, 2015; Novy-Marx, 2013), net payout yield (Boudoukh et al., 2007), momentum (Jegadeesh and Titman, 1993; Carhart, 1997), intermediate horizon return (Novy-Marx, 2012), the lagged monthly return to capture return reversal (Jegadeesh, 1990), price impact (Amihud, 2002), turnover (CHS), and the maximum return measure of Bali et al. (2011) (the average of the five highest daily returns within a month).[15] The construction of all these measures is detailed in Appendix A.

Our methodology differs from BMV and CHS in that we use cross-sectional ranks of predictors on the right-hand side of Eq. (4), and we forecast the cross-sectional rank of the risk measure on the left-hand side. Our methodology further differs from BMV because we use all past observations instead of estimating cross-sectional regressions separately for each month. In addition to predicting only the information needed to form the coskewness factor, another advantage of our approach is that computing cross-section ranks ensures we use stationary variables and the regression coefficients are not impacted by outliers possibly caused by data errors. Given that all regressors are transformed into cross-sectional ranks, our method also allows us to easily compare coefficient values across risk measures and firm characteristics.

2.4. Performance of the predicted systematic skewness factor

We next present in the seventh row of Table 1 (see row label PSS) the summary performance statistics of the PSS factor. The factor has a positive and significant average excess return of 5.37% per year and a Sharpe ratio of 0.38, almost double the maximum Sharpe ratio of the other coskewness factors. Most importantly, it has a significantly negative realized coskewness of −0.27 compared to a value of 0 for the market portfolio. Although we only report the realized monthly standardized coskewness

[13] Given that we estimate linear regressions, the predicted cross-sectional ranks can fall outside of the 0 to 1 interval. As we use predicted ranks only to classify stocks into high- and low-predicted coskewness portfolios, this does not affect our methodology.

[14] Inference about the panel regression coefficients is complicated by two aspects: the period used to compute the risk measures overlap and we estimate cross-sectional ranks with error. We avoid these complications because we rely only on the predicted values $F(\widehat{CS_{t\to t+11,i}})$ to form a long-short portfolio and do not make any inference about the regression coefficients.

[15] In a robustness check, we also used industry dummies. Results are available from the author.

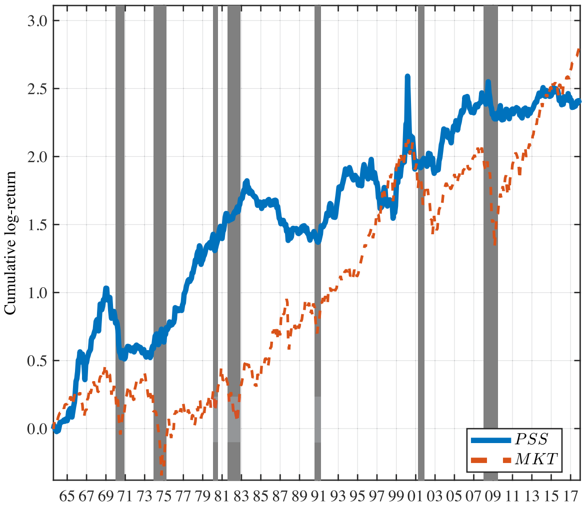

Fig. 1. Cumulative log-return for the *PSS* and *MKT* factors. We report the cumulative log-return of 1$ invested in the predicted systematic skewness factor *PSS* and the excess return on the market portfolio *MKT* from July 1963 to December 2017. Gray areas denote NBER recessions. Both factors are self-financed.

measure β_{HS}, we also compute the daily standardized coskewness, the daily and monthly realized coskewness *Cos*, and the daily and monthly realized β_{M^2}. In all cases, we find that the *PSS* factor creates negative realized coskewness measures that are the lowest across all factors.

When judged against factor models, we find an annualized α of 4.02% for the CAPM, −0.36% for the four-factor model, and 5.91% for the five-factor model, the first and last of which are significant. While the *PSS* factor adds the least amount of information when the momentum factor is included in these regressions, we show in Section 3.2 that *PSS* helps in pricing test portfolios even in the presence of *MOM*. Therefore, the *PSS* factor is successful in creating an ex post spread in coskewness, has a significant and positive average excess return, and adds value to other leading risk factors used in the literature.[16]

Fig. 1 reports the cumulative log-return of the *PSS* and *MKT* factors along with gray bars for the National Bureau of Economic Research (NBER) recessions. The *PSS* factor crashes before the 1970 recession, at the onset of the Asian crisis in 1997, before the 2001 recession, and during the financial crisis in 2008. On the other hand, the *PSS* factor does not go down as the market plummets during the 1973–1975 and 1981–1982 recessions. This behavior is in line with the low market β_M of 0.21 reported in the fifth column of Table 1. Fig. 1 also shows that the positive performance of the *PSS* cannot be attributed to a specific period. Overall, we find that the *PSS* factor is distinct from the market factor *MKT*.

[16] In unreported results, we find that the significance of the αs are unchanged when we augment the four-factor and five-factor models with the illiquidity factor of Pastor and Stambaugh (2003).

2.5. Why does our methodology work?

There are two main reasons why our methodology produces better results than using rolling coskewness estimates. First, we are using more information from a longer estimation sample and multiple predictors. Second, by working on cross-sectional ranks, we are attenuating the impact of time-varying marginal distributions on the estimation of regression coefficients in Eq. (4). We explain each reason below.

2.5.1. Using more information

Sorting stocks based on a lagged measure of coskewness, using either the past 60 monthly returns or one year of daily returns, is bound to have limited predictive ability for future coskewness. To distinguish how different stocks react to extreme events, one needs an estimation sample during which such events have occurred (or at least large values of r^2_M). When no such event has occurred during the estimation period, past coskewness has a limited ability to predict future coskewness.

In contrast, our panel approach considers all past observations and uses lagged firm characteristics and other risk measures in addition to lagged coskewness. Using firm characteristics helps capture, on a timely basis, the changing nature of individual stocks. Consider an example in which coskewness is positively related to firm size (large firms have higher coskewness than small firms). When a firm goes from being a small firm to a large firm, a rolling estimate of coskewness is not only dependent on the presence of extreme events in the lookback period, but it would also lag the true higher coskewness level of the newly large firm. In contrast, using the firm size would readily capture its new coskewness level.

2.5.2. Using cross-sectional ranks

In this section, we motivate the use of cross-sectional ranks for predictors and the predicted variable. For expositional simplicity, consider the predictive cross-sectional regression $Y_t = X_{t-1}\beta + \varepsilon_t$ where Y_t is a N-by-one vector, X_{t-1} is a N-by-K matrix of predictors, and ε_t is a N-by-one vector of random shocks.

If we want to form a portfolio whose composition rely on a quantile (e.g., the bottom 30% or top 30%), then our objective is to maximize the predictive Spearman rank correlation between the realized values, Y_{T+1}, and our predictions, $\hat{Y}_{T+1} = X_T\hat{\beta}$,

$$\rho^S\left(Y_{T+1}, \hat{Y}_{T+1}\right) = 12\left(E\left[G(Y_{T+1})H\left(\hat{Y}_{T+1}\right)\right] - \frac{1}{4}\right). \tag{6}$$

In this equation, $G(\cdot)$ is the cross-sectional CDF of Y_{T+1}; it governs how Y_{T+1} is distributed across the N stocks at a point in time, not how it is distributed over time for a given stock. Similarly, $H(\cdot)$ is the cross-sectional CDF of $\hat{Y}_{T+1}$. The CDF of a random variable always has a uniform distribution between 0 and 1 with an expected value of $\frac{1}{2}$ and a variance of $\frac{1}{12}$. Therefore, the predictive rank correlation in Eq. (6) only depends on the expected product of $G(Y_{T+1})$ and $H(\hat{Y}_{T+1})$.

In a finite sample, this predictive rank correlation is at its maximum value of one when the normalized ranks—the sample estimator of the CDF—are the same for $\hat{Y}_{T+1}$ as for Y_{T+1}. In such case, the composition of a portfolio based on a quantile of $\hat{Y}_{T+1}$, for example using the bottom 30% of stocks sorted by coskewness, will be the same as the one based on the quantile of Y_{T+1}. Conversely, if the rank correlation is equal to -1 then the low-$\hat{Y}_{T+1}$ portfolio will contain the stocks with the highest Y_{T+1} and the high-$\hat{Y}_{T+1}$ portfolio will contain the lowest Y_{T+1} stocks.

Why would using cross-sectional ranks lead to a higher predictive rank correlation? Computing cross-sectional ranks attenuates the impact of time-variations in the cross-sectional marginal distributions of the Xs and Y on the estimated regression coefficients β. Given that rank correlations only depend on the underlying dependence function through $E[G(Y_{T+1})H(\hat{Y}_{T+1})]$, working with cross-sectional ranks avoids having the estimated β being polluted by time-varying marginal distributions.[17]

We provide the intuition for this result using a simple example in which we have only one cross-section to estimate β ($T=2$) and Y_t and X_{t-1} are jointly normally distributed each period. Further, we assume the correlation matrix of X_{t-1} and Y_t is constant over time. In Section 5, we provide a comprehensive Monte Carlo simulation with multiple time periods, non-normal distributions, and measurement errors.

When regressing Y_2 on X_1, we obtain the OLS beta as $\hat{\beta}=\Sigma_{x,x,1}^{-1}\Sigma_{x,y,1}$, where $\Sigma_{x,x,1}$ is the K-by-K covariance matrix of X_1, and $\Sigma_{x,y,1}$ is the K-by-1 vector of covariances between X_1 and Y_2.

The Spearman rank correlation, ρ^S, of two jointly normally distributed variables is a monotonically increasing function of its Pearson linear correlation, ρ, as $\rho^S=\frac{6}{\pi}sin^{-1}(\frac{\rho}{2})$. Therefore, maximizing the predictive rank correlation in this example is equivalent to maximizing the predictive linear correlation between predicted and realized values of Y_{t+1}.

The linear correlation between the predicted values, $\hat{Y}_3=X_2\hat{\beta}$, and the realized values, Y_3, is

$$\frac{\Sigma_{x,y,2}^{\top}\Sigma_{x,x,1}^{-1}\Sigma_{x,y,1}}{\sqrt{\Sigma_{y,2}\left(\Sigma_{x,y,1}^{\top}\Sigma_{x,x,1}^{-1}\Sigma_{x,x,2}\Sigma_{x,x,1}^{-1}\Sigma_{x,y,1}\right)}}, \tag{7}$$

where $\Sigma_{y,2}$ is the variance of Y_3.

The highest predictive correlation we could obtain is if the true $\beta=\Sigma_{x,x,2}^{-1}\Sigma_{x,y,2}$ that links X_2 to Y_3 was known. In this case, the correlation would be

$$\sqrt{\frac{\Sigma_{x,y,2}^{\top}\Sigma_{x,x,2}^{-1}\Sigma_{x,y,2}}{\Sigma_{y,2}}}. \tag{8}$$

Given that Eq. (8) is the highest predictive correlation, the one based on $\hat{\beta}$ in Eq. (7) is therefore lower whenever $\Sigma_{x,x,1}\neq\Sigma_{x,x,2}$, $\Sigma_{x,y,1}\neq\Sigma_{x,y,2}$, or both. As we have assumed correlations between variables to be constant, it is the time variation in their variances that causes the loss in predictive correlation.[18] In contrast, working on a joint distribution where the marginal distributions are constant over time—as one does by working with normalized ranks—avoids this problem because all marginal distribution parameters cancel out in Eq. (7).[19] The estimated regression coefficients are $\hat{\beta}_F=C_{x,x}^{S,-1}C_{x,y}^{S}$, where $C_{x,x}^{S}$ is the K-by-K matrix with rank correlations of X_1, and $C_{x,y}^{S}$ is the K-by-1 vector of rank correlations between X_1 and Y_2. These rank correlations do not have time subscripts because we have assumed correlations to be constant over time. The predictive rank correlation

$$\sqrt{\frac{C_{x,y}^{S,\top}C_{x,x}^{S,-1}C_{x,y}^{S}}{C_{y}^{S}}} \tag{9}$$

is not impacted by marginal distributions, in contrast with the predictive correlation in Eq. (7).

There are two cases in which our methodology is not necessary. First, if you have one predictor and X and Y are jointly normally distributed, then variances cancel out in Eq. (7), and the OLS coefficient is not be affected by time-varying variances. We show in the next section that multiple variables are important predictors of future coskewness, and therefore this is not an issue.

The second case in which our methodology does not improve forecasts is if the marginal distributions of the Xs and Y do not vary over time. In the Online Appendix, we report the time series of the cross-sectional quantiles of coskewness in Fig. 1 and of the cross-sectional quantiles of all predictors in Fig. 2. Fig. 1 shows that the cross-sectional distribution of coskewness is drastically different during periods surrounding the 1987 stock market crash, the dot-com bubble burst, and the financial crisis of 2008–2009, compared to other periods. Similarly, Fig. 2 shows important time variations in the cross-sectional quantiles of most predictors. Therefore, the marginal distributions of the Xs and Y are not constant over time.

In the normally distributed case, another possibility would be to simply standardize (i.e., remove the mean and divide by the standard deviation) all variables to neutralize the impact of time-varying marginal distributions. However, standardizing neutralizes the impact of marginal distributions only when those are normal. In contrast, computing normalized ranks also neutralizes time-varying higher order moments in the cross-sectional marginal distributions.

The limit of our argument is that there could be time variations not only in the marginal distributions but also in the underlying dependence function. To address this limitation, we could estimate a time-varying dependence model but this would considerably complicate our approach relative to running simple linear regressions. As a simpler solution, we estimate Eq. (4) using ten-year rolling

[17] We denote as dependence function the copula between the Xs and Y. A copula is a joint distribution function in which all variables have a uniform marginal distribution and denotes the part of a joint distribution before random variables are mapped to their respective marginal distributions.

[18] More generally, in large T cases, the same intuition holds as long as the covariance matrix of X_T and/or the covariances between X_T and Y_{T+1} are different from their unconditional counterparts.

[19] The regression coefficients are still impacted by the errors made when computing normalized ranks. Simulation evidence in Section 5 shows this is not an issue with the large cross-section of stocks we use.

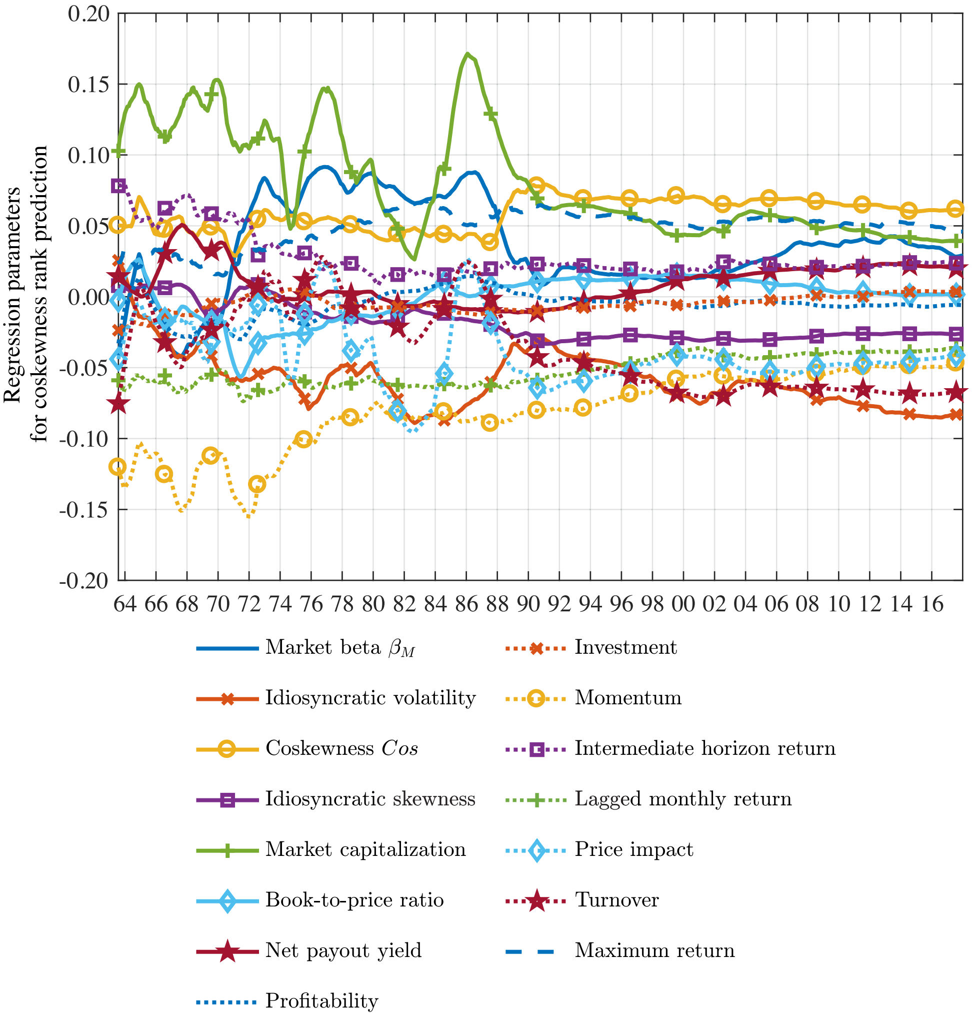

Fig. 2. Coefficients of predictive panel regressions for coskewness ranks. We report the panel regression coefficients $\hat{\theta}$ and $\hat{\phi}$ in Eq. (4) from July 1963 to December 2017. Each month, we run a panel regression that predicts the next 12-month realized daily coskewness using past risk measures and stock characteristics. We use the normalized cross-sectional rank of coskewness as the dependent variable and the normalized cross-sectional ranks of past risk measures and characteristics as predictors. We use all past observations to estimate at each point in time the panel regression. The construction of all variables is detailed in Appendix A.

windows. The results (unreported) are not as good, suggesting that there are no significant time variations in how future coskewness is related to past risk measures and firm characteristics.

We also run other robustness checks. We consider a forecast combination method by estimating Eq. (4) using only a constant and one predictor at a time and then computing the average forecast from all models (see, for example, Rapach et al., 2010). Using either all past data or rolling estimation windows, the results are not as good, indicating that our methodology benefits from cross-correlations between predictors to obtain better coskewness rank predictions.

To further illustrate the importance of using cross-sectional ranks in our panel regression, the next line in Table 1 reports the results for a long-short factor built using the raw (i.e., unranked) coskewness values on the left-hand side of regression (4). The average return of 2.52 is lower and insignificant; the Sharpe ratio of 0.19 is half the 0.38 value for the *PSS* factor; the factor regression α is significant for the five-factor model, but smaller in magnitude; and the realized standardized coskewness measure, though significant, is less negative (−0.16 compared to −0.27). We also build a factor using the unranked values of both the predictors on the right-hand side and coskewness on the left-hand side (unreported results). This factor performs even worse: it has a realized coskewness close to zero, and all its factor regression αs are insignificant. In the next section, we also discuss that the panel regression coefficients θ and ϕ for these two factors are unstable, whereas we find stable coefficients when using cross-sectional ranks of coskewness and the predictors. Therefore, we better capture ex ante the coskewness risk factor by predicting cross-sectional ranks using the cross-sectional ranks of predictors than by using unranked variables.

2.6. What predicts systematic skewness?

In this section, we describe the predictive power of each risk measure and firm characteristics. Fig. 2 reports the panel regression coefficients $\hat{\theta}$ and $\hat{\phi}$ in Eq. (4) estimated over time using expanding samples, while Table 1 of the Online Appendix provides the time-series averages of coefficients as well as their 5th and 95th percentiles.

Working with cross-sectional ranks results in stable regression estimates, especially in the later part of the sample period when the estimation sample is longer.[20] Lagged coskewness, market capitalization, intermediate horizon return, and maximum return always positively predict coskewness. Idiosyncratic volatility, momentum, and lagged monthly return always negatively predict coskewness, although idiosyncratic volatility's coefficient is briefly positive at the start of the sample period.

CHS shows that returns over the last six months negatively predict daily total return skewness. Harvey and Siddique (2000) find a strong negative relation between return momentum (measured from month $t-12$ to $t-2$) and coskewness. Novy-Marx (2012) show that the momentum effect mainly comes from intermediate horizon returns (from month $t-12$ to $t-7$), not from recent returns. The coefficients for momentum and lagged monthly return (return for month $t-1$) indicate that they robustly predict lower coskewness, whereas intermediate horizon return predicts higher coskewness. Therefore, returns over the last six months have more importance in predicting lower coskewness, similar to CHS's results for total skewness.

We find that market capitalization is positively related to coskewness; large firms have higher coskewness than small firms. In contrast, Conrad et al. (2013) use option-implied moments over the 1996–2005 period to compute a risk-neutral estimate of coskewness and find that it is negatively related to firm size. CHS and BMV find that large firms have more negative daily total skewness and monthly idiosyncratic skewness, respectively. In Section 4, we use a predictive panel regression for idiosyncratic skewness to distinguish the determinants of systematic from those of idiosyncratic skewness. We find that large firms have lower idiosyncratic skewness, in line with their results. We further show that predictors for coskewness differ from those for idiosyncratic skewness, which demonstrates that the coskewness factor is a distinct source of risk.

The time-series coefficient averages reported in Table 1 of the Online Appendix shed light on the superior performance of the *PSS* factor in Table 1 compared to factors based only on past coskewness measures. The coefficient θ for lagged coskewness is always positive, indicating that coskewness is persistent. But its average value of 0.059 is small, and past coskewness is only one of the different predictors. Hence, the *PSS* factor better captures future coskewness risk by using more information.

Having established that the *PSS* factor successively captures future realized coskewness, we proceed in the next section to formal asset pricing tests.

3. Systematic skewness and expected stock returns

We explore the pricing performance of the *PSS* factor in conjunction to leading factor models. We use three sets of test portfolios of US equities: 25 size and book-to-market ratio sorted portfolios, 25 size and momentum sorted portfolios, and 25 size and predicted coskewness sorted portfolios.[21] Following the advice of Lewellen et al. (2010), we report the GLS R^2 and include a constant to allow for a zero-beta rate different from the risk-free rate. Further, we use the methodology of Kan et al. (2013) to compute model misspecification robust t-ratios of risk premia estimates, prices of risk, and standard errors of the cross-sectional R^2.

For each of the $N=25$ test portfolios, we run time-series regressions of their excess returns on a constant and factor returns. Then, we run a cross-sectional regression of the test portfolio average returns on their estimated time-series factor loadings to estimate the factor risk premia. We estimate different factor models to investigate the added pricing power of the *PSS* factor. We start in the next section with the CAPM.

3.1. Test of the CAPM

We start in this section with a comparison of the CAPM and the CAPM augmented with the *PSS* factor. Table 2 reports estimation results for the 25 size and book-to-price portfolios in Panel A, the 25 size and momentum portfolios in Panel B, and the 25 size and coskewness portfolios in Panel C.

For all models, we report the model misspecification robust t-ratios in parentheses below the estimates of the constant and risk premia. We then report the R^2 and p-value in parenthesis for the increase in R^2 for each model compared to the one above. All parameters are estimated by GLS, but we report both OLS and GLS R^2.

In both panels, we first report the pricing performance for the CAPM. The risk premium is negative, in line with previous results (see, for example, Kan et al., 2013). When we add the *PSS* factor, we obtain positive risk premia of 0.64%, 0.74%, and 0.47% per month for the 25 size and book-to-price ratio, the 25 size and momentum, and the 25 size and coskewness portfolios, respectively. The corresponding model misspecification robust t-ratios are 2.62, 3.06, and 2.94.

The increases in model fit in Panel A and B are striking. The OLS R^2s increase from 0.08 to 0.51 for size and book-to-market portfolios and from 0.11 to 0.77 for size and momentum portfolios. The GLS R^2s increase from 0.11 to 0.29 and from 0.02 to 0.20, respectively. All R^2 increases are significant. For the 25 size and coskewness portfolio, the OLS

[20] In unreported results, we found that using unranked variables resulted in unstable regression coefficient values.

[21] We obtain the size, book-to-market, and momentum portfolios from Kenneth French's website. We build the 25 size and coskewness portfolios by first sorting each month on market capitalization and then on predicted coskewness ranks.

Table 2
Is the *PSS* factor significant?
We report asset pricing tests for the CAPM with the market excess return (*MKT*) and the CAPM augmented with the predicted systematic skewness factor (*PSS*). We also report on a model where the *MKT* factor is separated into low returns (MKT^-) and high returns (MKT^+). We define low returns as those below the average market excess return minus one standard deviation. As test assets, we use 25 size and book-to-price ratio sorted US equity portfolios in Panel A, 25 size and momentum sorted US equity portfolios in Panel B, and 25 size and predicted coskewness sorted US equity portfolios in Panel C. For each test portfolio, we run a time-series regression of its excess returns on a constant and factor returns. Then, we run a cross-sectional regression of average test portfolio excess returns on a constant and betas obtained from the time-series regressions. For each model, we estimate the beta form and report the risk premia in % per month for each factor in the first two rows. We then estimate the model in its covariance form and report the prices of risk in the last two rows for each model. We report the model misspecification robust *t*-ratios from Kan et al. (2013) below risk premia and prices of risk. Below the R^2s for the "*MKT, PSS*" model, we report the *p*-value for the one-sided test that the model has a significantly higher R^2 than the CAPM. Below the R^2s for the last model, we report the *p*-value for the two-sided test for non-nested models that the model has a significantly different R^2 than the "*MKT, PSS*" model (see Kan et al., 2013). The data are monthly from July 1963 to December 2017.

Model		*Constant*	*MKT*	*PSS*	MKT^-	MKT^+	*OLS* R^2	*GLS* R^2
Panel A: 25 size and book-to-price ratio portfolios								
CAPM	Risk premium	1.36	−0.80				0.08	0.11
		(5.27)	(−2.62)					
	Price of risk	0.01	−4.16					
		(5.27)	(−2.65)					
MKT, PSS	Risk premium	1.71	−1.15	0.64			0.51	0.29
		(6.08)	(−3.64)	(2.62)			(0.04)	(0.01)
	Price of risk	0.02	−7.13	5.55				
		(6.08)	(−3.77)	(2.65)				
PSS, β_M^-, β_M^+	Risk premium	1.68		0.62	−0.79	−0.34	0.52	0.30
		(6.00)		(2.47)	(−1.41)	(−0.56)	(0.66)	(0.76)
	Price of risk	0.02		5.78	−12.35	−2.69		
		(6.00)		(2.70)	(−1.29)	(−0.33)		
Panel B: 25 size and momentum portfolios								
CAPM	Risk premium	0.99	−0.36				0.11	0.02
		(4.14)	(−1.26)					
	Price of risk	0.01	−1.85					
		(4.14)	(−1.27)					
MKT, PSS	Risk premium	1.41	−0.77	0.74			0.77	0.20
		(4.45)	(−2.29)	(3.06)			(0.00)	(0.00)
	Price of risk	0.01	−5.21	5.69				
		(4.45)	(−2.55)	(2.91)				
PSS, β_M^-, β_M^+	Risk premium	1.61		0.82	−0.93	−0.06	0.77	0.22
		(2.87)		(2.64)	(-0.72)	(-0.08)	(1.00)	(0.65)
	Price of risk	0.02		7.13	−15.52	0.67		
		(2.87)		(2.02)	(−0.75)	(0.06)		
Panel C: 25 size and coskewness portfolios								
CAPM	Risk premium	−0.19	0.80				0.95	0.11
		(−0.36)	(1.32)					
	Price of risk	−0.00	4.18					
		(−0.36)	(1.29)					
MKT, PSS	Risk premium	0.47	0.11	0.47			0.98	0.30
		(0.77)	(0.16)	(2.94)			(0.08)	(0.02)
	Price of risk	0.00	−0.03	2.81				
		(0.77)	(−0.01)	(2.31)				
PSS, β_M^-, β_M^+	Risk premium	0.42		0.47	−0.06	0.21	0.98	0.31
		(0.67)		(2.96)	(−0.13)	(0.31)	(1.00)	(0.84)
	Price of risk	0.00		2.93	−2.30	2.25		
		(0.67)		(2.53)	(−0.32)	(0.27)		

(GLS) R^2 increases from 0.95 (0.11) to 0.98 (0.30). Both increases are significant at the 10% level.

To judge the added value of a risk factor, Kan et al. (2013) stress that it is the *t*-ratio of its price of risk, not its risk premium, that should be compared to a critical value. Accordingly, we report in the last two rows for each model the price of risk estimates obtained by estimating the covariance form of the asset pricing model instead of the beta form. We find in Panel A that the *PSS* factor carries a price of risk of 5.55 with a *t*-ratio of 2.65. The corresponding values in Panel B and C are 5.69 with a *t*-ratio of 2.91 and 2.81 with a *t*-ratio of 2.31.

We next compare the *PSS* pricing performance to the downside β_M measure of Ang et al. (2006) and Lettau et al. (2014). To capture downside risk, they measure a portfolio's sensitivity to market returns when it is below a threshold. We follow Lettau et al. (2014) and define the threshold as the average excess market return minus one standard deviation.[22] In the first step time-series regression, we regress returns on the *PSS*, a MKT^- factor that is equal to $r_{M,t}$ if $r_{M,t}$ is below the threshold and zero

[22] Over the July 1963 to December 2017 sample period, 90 of the 654 market returns are below this threshold.

Table 3
Is the *PSS* factor significant in the four-factor model?
We report asset pricing tests for the four-factor model with market (*MKT*), size (*SMB*), value (*HML*), and momentum (*MOM*) factors and a four-factor model with *MKT, PSS, HML*, and *MOM*. As test assets, we use 25 size and book-to-price ratio sorted US equity portfolios in Panel A, 25 size and momentum sorted US equity portfolios in Panel B, and 25 size and predicted coskewness sorted US equity portfolios in Panel C. For each test portfolio, we run a time-series regression of its excess returns on a constant and factor returns. Then, we run a cross-sectional regression of average test portfolio excess returns on a constant and betas obtained from the time-series regressions. For each model, we estimate the beta form and report the risk premia in % per month for each factor in the first two rows. We then estimate the model in its covariance form and report the prices of risk in the last two rows for each model. We report the model misspecification robust *t*-ratios from Kan et al. (2013) below risk premia and prices of risk. Below the R^2s for the last model, we report the *p*-value for the two-sided test for non-nested models that the model has a significantly different R^2 than the four-factor model (see Kan et al., 2013). The data are monthly from July 1963 to December 2017.

Model		*Constant*	*MKT*	*PSS*	*SMB*	*HML*	*MOM*	*OLS* R^2	*GLS* R^2
Panel A: 25 size and book-to-price ratio portfolios									
Four-factor	Risk premium	1.07	−0.50		0.22	0.35	1.57	0.76	0.40
		(2.50)	(−1.12)		(1.72)	(2.64)	(1.47)		
	Price of risk	0.01	−0.96		4.05	7.52	9.74		
		(2.50)	(−0.25)		(2.01)	(2.09)	(1.38)		
MKT, PSS, HML, MOM	Risk premium	1.18	−0.60	0.98		0.35	1.59	0.76	0.42
		(2.69)	(−1.34)	(2.39)		(2.65)	(1.51)	(0.93)	(0.65)
	Price of risk	0.01	−1.87	4.48		7.07	8.29		
		(2.69)	(−0.49)	(1.95)		(1.98)	(1.18)		
Panel B: 25 size and momentum portfolios									
Four-factor	Risk premium	1.06	−0.44		0.33	0.02	0.69	0.85	0.26
		(1.75)	(−0.76)		(2.25)	(0.04)	(3.80)		
	Price of risk	0.01	−2.59		4.87	1.20	3.71		
		(1.75)	(−0.56)		(2.11)	(0.15)	(1.66)		
MKT, PSS, HML, MOM	Risk premium	0.98	−0.36	0.66		0.20	0.70	0.86	0.24
		(1.64)	(−0.64)	(2.61)		(0.50)	(3.89)	(0.99)	(0.51)
	Price of risk	0.01	−1.72	3.79		3.63	2.99		
		(1.64)	(−0.39)	(1.77)		(0.54)	(1.33)		
Panel C: 25 size and coskewness portfolios									
Four-factor	Risk premium	0.69	−0.12		0.53	0.17	−0.33	0.98	0.40
		(0.97)	(−0.15)		(3.04)	(0.50)	(−0.90)		
	Price of risk	0.01	−1.85		6.79	2.28	−1.83		
		(0.97)	(−0.42)		(2.46)	(0.59)	(−0.85)		
MKT, PSS, HML, MOM	Risk premium	1.01	−0.45	0.48		0.18	−0.51	0.98	0.43
		(1.30)	(−0.53)	(2.99)		(0.56)	(−1.21)	(0.88)	(0.41)
	Price of risk	0.01	−4.04	5.55		0.74	−5.06		
		(1.30)	(−0.81)	(2.43)		(0.21)	(−1.57)		

otherwise, and a MKT^+ factor that is equal to $r_{M,t}$ if $r_{M,t}$ is above the threshold and zero otherwise. To capture a downside risk premium, we should expect the risk premium for MKT^- to be higher than the premium for MKT^+.

We find for both sets of test portfolios that the *PSS* risk premium and prices of risk remain significant, whereas coefficients for both MKT^- and MKT^+ are not significant. Overall, we conclude that the *PSS* factor carries a significant amount of information in pricing size, book-to-market, momentum, and coskewness sorted portfolios.

3.2. Test of multi-factor models

In this section, we examine whether the *PSS* factor adds any pricing information in multi-factor models. We report on the four-factor model with *MKT, SMB, HML*, and *MOM* in Table 3 and on the five-factor model with *MKT*, SMB_{FF5}, *HML, RMW*, and *CMA* in Table 4. These tables have the same structure as Table 2.

In the first two rows of each panel in Table 3, we report on the estimates for the four-factor model. *HML* is only significant for size and book-to-market portfolios and *MOM* is only significant for size and momentum portfolios, whereas *SMB* is significant for all sets of portfolios. The OLS R^2 are, respectively, 0.76, 0.85, and 0.98, and the GLS R^2 are 0.40, 0.26, and 0.40.

When augmenting the four-factor model with *PSS* (unreported results), we initially find that the prices of risk for *PSS* and *SMB* are not significant, and the R^2 is not significantly higher than for the four-factor model. This result can be understood as follows. *SMB* and *PSS* have a high correlation of 0.73 during the 1963–2017 period.[23] We explore the interactions between *SMB* and *PSS* in Section 3.4 below. Despite their high correlation, we find that *PSS* contains more pricing information than *SMB*. Therefore, we remove *SMB* from all factor models.

The last four lines in each panel report on the models with *MKT, PSS, HML*, and *MOM*. All coskewness risk premium and price of risk estimates are positive and significant at the 5% level. We further find that the R^2 is not significantly different from the four-factor model R^2 (neither lower nor higher).[24] Hence, we replace the *SMB*

[23] The size factor in the Fama–French five-factor model examined below is constructed differently. That size factor has a correlation of 0.71 with *PSS*.

[24] We use the normal test for non-nested models of Kan et al. (2013).

Table 4
Is the *PSS* factor significant in the five-factor model?
We report asset pricing tests for the five-factor model with market (*MKT*), size (SMB_{FF5}), value (*HML*), profitability (*RMW*), and investment (*CMA*) factors and a five-factor model with *MKT, PSS, HML, RMW*, and *CMA*. As test assets, we use 25 size and book-to-price ratio sorted US equity portfolios in Panel A, 25 size and momentum sorted US equity portfolios in Panel B, and 25 size and predicted coskewness sorted US equity portfolios in Panel C. For each test portfolio, we run a time-series regression of its excess returns on a constant and factor returns. Then, we run a cross-sectional regression of average test portfolio excess returns on a constant and betas obtained from the time-series regressions. For each model, we estimate the beta form and report the risk premia in % per month for each factor in the first two rows. We then estimate the model in its covariance form and report the prices of risk in the last two rows for each model. We report the model misspecification robust t-ratios from Kan et al. (2013) below risk premia and prices of risk. Below the R^2s for the last model, we report the p-value for the two-sided test for non-nested models that the model has a significantly different R^2 than the five-factor model (see Kan et al., 2013). The data are monthly from July 1963 to December 2017.

Model		*Constant*	*MKT*	*PSS*	SMB_{FF5}	*HML*	*RMW*	*CMA*	*OLS* R^2	*GLS* R^2
Panel A: 25 size and book-to-price ratio portfolios										
Five-factor	Risk premium	1.37	−0.83		0.26	0.34	0.23	0.19	0.78	0.32
		(4.18)	(−2.31)		(2.04)	(2.57)	(1.03)	(0.79)		
	Price of risk	0.01	−4.48		6.05	3.20	5.24	−1.07		
		(4.18)	(−1.66)		(3.34)	(0.56)	(1.10)	(−0.09)		
MKT, PSS, HML, RMW, CMA	Risk premium	1.32	−0.77	0.80		0.34	0.33	0.20	0.84	0.48
		(4.35)	(−2.29)	(3.01)		(2.57)	(1.86)	(1.00)	(0.12)	(0.00)
	Price of risk	0.01	−3.39	9.57		3.06	13.10	3.39		
		(4.35)	(−1.34)	(4.39)		(0.62)	(2.72)	(0.33)		
Panel B: 25 size and momentum portfolios										
Five-factor	Risk premium	1.12	−0.49		0.44	−0.63	0.25	0.52	0.89	0.43
		(2.04)	(−0.90)		(3.27)	(−1.66)	(0.69)	(1.27)		
	Price of risk	0.01	−0.00		9.17	−29.77	13.22	43.82		
		(2.04)	(−0.00)		(2.70)	(−2.48)	(1.39)	(2.21)		
MKT, PSS, HML, RMW, CMA	Risk premium	0.89	−0.26	0.59		−0.10	0.02	0.75	0.90	0.43
		(1.72)	(−0.52)	(2.33)		(−0.34)	(0.06)	(2.06)	(0.75)	(0.82)
	Price of risk	0.01	2.73	7.27		−21.22	11.00	44.92		
		(1.72)	(0.68)	(2.40)		(−1.95)	(1.04)	(2.37)		
Panel C: 25 size and coskewness portfolios										
Five-factor	Risk premium	0.74	−0.15		0.47	0.07	−0.05	−0.23	0.98	0.47
		(0.99)	(−0.18)		(2.52)	(0.15)	(−0.21)	(−0.80)		
	Price of risk	0.01	−3.33		5.70	7.46	−0.89	−14.97		
		(0.99)	(−0.67)		(1.88)	(0.81)	(−0.12)	(−1.23)		
MKT, PSS, HML, RMW, CMA	Risk premium	0.54	0.05	0.47		0.02	−0.10	−0.22	0.98	0.35
		(0.75)	(0.06)	(2.89)		(0.04)	(−0.35)	(−0.81)	(0.96)	(0.00)
	Price of risk	0.01	−1.51	2.37		6.32	−1.64	−12.06		
		(0.75)	(−0.30)	(1.06)		(0.54)	(−0.16)	(−0.77)		

factor with the theoretically motivated *PSS* factor without impacting the model fit.

Finally, we examine the five-factor model in Table 4. For size and book-to-market ratio test portfolios in Panel A, the *PSS* factor price of risk is positive and significant. The increases in GLS R^2 is significant at the 1% level, whereas the increase in OLS R^2 is marginally significant with a p-value of 0.12.

For size and momentum portfolios in Panel B, both risk premium of 0.59% per month and price of risk of 7.27 are significant. The p-values for model comparison tests reported in parenthesis below the R^2 indicate that they are not significantly different. Hence, we replace SMB_{FF5} with the theoretically motivated *PSS* without impacting the model fit.

For size and coskewness portfolios in Panel C, the risk premium is 0.47% per month with a t-ratio of 2.89. The price of risk is positive but not significant. In unreported results, we find that removing any one of the *HML, RMW*, or *CMA* factors leads to a significant price of risk for *PSS*.

Overall, our results show that measuring ex ante systematic skewness is important for explaining the cross-section of stock returns, even in the presence of factors built to explain the cross-section of average returns of the chosen test portfolios (e.g., *MOM* for momentum sorted portfolios).

3.3. Robustness check

As a robustness check, we use the methodology of Giglio and Xiu (2017) that accounts for omitted factors and measurement error when estimating a factor risk premium. Table 5 contains the estimated constant and risk premia for the *MKT* and *PSS* factors.[25] Their methodology first obtains the first p principal components from test portfolio returns, computes the loadings by regressing portfolio returns on the principal components, and then obtains the principal components' risk premia from a cross-sectional regression of average test portfolio returns on their loadings. Then, the risk premia of the factors of interest are obtained by combining the principal components' risk premia and the time-series regression loadings of risk factors on principal components.

Their methodology relies on a large cross-section of test portfolios to extract a set of principal components.

[25] We thank Stefano Giglio for making their estimation code available on his website.

Table 5
Robustness check—asset pricing tests with omitted factors.
We report estimated constants and risk premia for the CAPM in which the market excess return (*MKT*) is augmented with the predicted systematic skewness factor (*PSS*). We use the estimation methodology of Giglio and Xiu (2017) that is robust to omitted factors and measurement error. As test assets, we use 25 size and book-to-price ratio, 25 size and momentum, and 25 size and predicted coskewness sorted US equity portfolios in the top rows. We use the 202 US equity portfolios from Giglio and Xiu (2017) in the bottom rows (25 portfolios sorted by size and book-to-market ratio, 17 industry portfolios, 25 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, 35 portfolios sorted by size and net issuance, 25 portfolios sorted by size and accruals, 25 portfolios sorted by size and momentum, and 25 portfolios sorted by size and beta). The second column reports on the number of principal components used to span the space of asset returns. The star denotes the optimal number as identified by their methodology. The last three columns report on the constant, the risk premium for *MKT*, and the risk premium for *PSS*, all reported in % per month. *t*-ratios are below risk premia in parentheses. The data are monthly from July 1963 to December 2017.

Test portfolios	p	*Constant*	*MKT*	*PSS*
75 size, book-to-price ratio, momentum, and coskewness portfolios	4	0.38	0.15	0.45
		(2.02)	(0.58)	(2.83)
	5	0.37	0.16	0.44
		(1.88)	(0.61)	(2.69)
	6*	0.46	0.07	0.44
		(1.73)	(0.23)	(2.74)
202 portfolios from Giglio and Xiu (2017)	4	0.21	0.35	0.28
		(2.26)	(1.73)	(1.89)
	5	0.26	0.29	0.30
		(2.52)	(1.38)	(2.00)
	6*	0.21	0.34	0.30
		(1.69)	(1.56)	(1.99)

Therefore, we combine the 25 size and book-to-market ratio portfolios, the 25 size and momentum portfolios, and the 25 size and coskewness portfolios. The second column in Table 5 reports the number of principal components p used in the tests. We follow Giglio and Xiu (2017) and report results using four, five, and six principal components. The optimal number identified by their methodology is six. The risk premium for *PSS* reported in the last column varies from 0.44% to 0.45% per month and is highly significant in all cases.

We also report model estimates using the set of 202 portfolios used in their paper that assembles the most well-known stock market anomalies. We find a significant risk premium for *PSS* ranging from 0.28% to 0.30%. While the premium based on such a large cross-section of portfolios is smaller, all estimates are significantly positive. The optimal number of principal components for this set of portfolios is six.[26]

3.4. Size and coskewness

In this section, we further explore the interactions between size and predicted coskewness. We report on double-sorted portfolio average returns and CAPM αs in Table 6 and factor spanning tests for *SMB* and *PSS* in Table 7.

Each month, we sort all available stocks into five quintiles based on their market capitalization at the end of the previous month. Within each size quintile, we form five value-weighted quintile portfolios based on predicted coskewness ranks.[27] Panel A of Table 6 reports average monthly returns for the 25 portfolios (columns Low *Cos* to High *Cos*) and the high-minus-low coskewnes portfolio (column High-Low) for each size quintile as well as the CAPM αs of the high-minus-low coskewness portfolios (last column). The first two lines contain the average returns and CAPM α of the quintile portfolios sorted only on predicted coskewness.

We find that there is a negative return spread between high and low coskewness portfolios within each size quintile. High-minus-low portfolios' average returns and CAPM αs are significant in all but one case (CAPM α for large stocks). Therefore, we find variations in average returns in coskewness portfolios even after controlling for the size effect.

Panel B reports on portfolios sorted first by predicted coskewness and then by market capitalization. While high-minus-low size portfolio average returns are negative in each coskewness quintile, they are significant only for the fourth and fifth coskewness quintile portfolios. However, none of the CAPM αs are significant. Therefore, we find little average return variations in size portfolios after controlling for coskewness. Results in Table 6 indicate that predictive coskewness contains more independent pricing information than size.

Another perspective on the relative importance of size and coskewness is presented in Table 7. We run time-series regression of *PSS*, the size factor *SMB* used in the four-factor model, and the size factor SMB_{FF5} used in the

[26] We run other robustness checks in the Online Appendix. The market risk premium estimates in Tables 2–5 are almost always insignificant. In Section 6 of the Online Appendix, we replicate these tables replacing *MKT* by a predicted covariance factor based on our predictive panel regression in Eq. (4) with future covariance on the left-hand side and find similar results. We provide in Section 7 of the Online Appendix a stylized three-moment CAPM showing why we should expect an expected return-to-β relation lower than the equity risk premium.

[27] We only explore conditional portfolio sorts. Given the positive correlation between market capitalization and predicted coskewness as reported in Section 2.6, forming 25 portfolios using an independent sort results in empty portfolios for some months.

Table 6
Size and predicted coskewness sorted portfolios.
We report average monthly returns (in %) of size and predicted coskewness conditionally sorted portfolios. Each month, we sort all available stocks into five quintiles based on their market capitalization at the end of the previous month. Within each size quintile, we form five quintile value-weighted portfolios based on predicted coskewness ranks. We report in Panel A the average return of each portfolio and of the high-minus-low portfolio as well as the CAPM α of the high-minus-low portfolio. In the first two lines of Panel A, we also report average returns and CAPM α for quintile portfolios sorted only on predicted coskewness ranks. Panel B reports on portfolios sorted by predicted coskewness rank first and then by size. We report in parentheses the *t*-ratios using a Newey-West estimator with $T^{0.25} \approx 6$ lags. The data are monthly from July 1963 to December 2017.

Panel A: size and predicted coskewness sorted portfolios

	Low *Cos*	2	3	4	High *Cos*	High-Low	High-Low CAPM α
	0.99	0.76	0.70	0.58	0.45	−0.54	−0.41
	(3.78)	(3.55)	(3.41)	(3.36)	(2.70)	(−2.86)	(−2.28)
Small	1.02	0.99	0.87	0.84	0.54	−0.49	−0.35
	(3.30)	(3.81)	(3.47)	(3.82)	(2.99)	(−2.23)	(−1.72)
2	1.03	0.94	0.77	0.68	0.49	−0.54	−0.41
	(3.72)	(3.80)	(3.58)	(3.26)	(2.91)	(−2.88)	(−2.30)
3	0.99	0.73	0.70	0.65	0.48	−0.51	−0.39
	(3.97)	(3.47)	(3.32)	(3.54)	(2.91)	(−3.12)	(−2.58)
4	0.81	0.68	0.59	0.51	0.45	−0.36	−0.28
	(3.80)	(3.40)	(3.06)	(2.96)	(2.72)	(−2.72)	(−2.14)
Large	0.70	0.57	0.49	0.45	0.46	−0.24	−0.19
	(3.47)	(3.24)	(2.85)	(2.71)	(2.55)	(−1.87)	(−1.46)

Panel B: predicted coskewness and size sorted portfolios

	Small	2	3	4	Large	High-Low	High-Low CAPM α
	0.87	0.84	0.79	0.71	0.48	−0.40	−0.24
	(2.95)	(3.36)	(3.66)	(3.71)	(2.89)	(−1.86)	(−1.22)
Low *Cos*	0.95	0.80	0.89	0.88	0.74	−0.21	−0.10
	(2.67)	(2.62)	(3.19)	(3.48)	(3.70)	(−0.93)	(−0.45)
2	0.91	0.86	0.81	0.82	0.64	−0.27	−0.15
	(2.85)	(3.17)	(3.33)	(3.72)	(3.49)	(−1.32)	(−0.78)
3	0.85	0.87	0.82	0.73	0.59	−0.26	−0.13
	(2.98)	(3.63)	(3.80)	(3.66)	(3.37)	(−1.43)	(−0.80)
4	0.85	0.79	0.71	0.64	0.51	−0.33	−0.21
	(3.39)	(3.64)	(3.58)	(3.41)	(3.00)	(−2.19)	(−1.42)
High *Cos*	0.79	0.71	0.55	0.53	0.45	−0.34	−0.21
	(3.61)	(3.78)	(2.97)	(3.04)	(2.75)	(−2.42)	(−1.55)

five-factor model on different factor models. For each regression, we report the monthly αs, factor exposures, and time-series R^2. While the *PSS* monthly αs reported in the first two lines of each Panel correspond to the annual αs reported in Table 1, we report them again for ease of comparison with the size αs.

We find that the size factors *SMB* and SMB_{FF5} have significant αs only when regressed against other factors in the five-factor model (see third and fifth regressions in Panel C). Their significant αs indicate that they add explanatory power when added to *MKT, HML, RMW*, and *CMA*.

However, none of their αs remain significant when we augment the model with *PSS* (see fourth and sixth regressions in Panel C). In contrast, *PSS*'s α remains significant after adding size to the model (see first and second regressions in Panel C). Therefore, while *PSS* and the size factors are highly correlated, we find that *PSS* contains more pricing information.

Now that we have shown the asset pricing importance of systematic skewness, we turn our attention to the relative importance of idiosyncratic skewness in the next section.

4. Idiosyncratic skewness

In this section, we investigate whether idiosyncratic skewness is priced once we account for systematic risk. In Section 4.1, we first predict the cross-sectional rank of idiosyncratic skewness and show that we capture differences in future realized idiosyncratic skewness across stocks. Then, we run factor analyses for predicted idiosyncratic skewness sorted portfolios in Section 4.2 and run some robustness checks in Section 4.3.

Models in Barberis and Huang (2008), Brunnermeier et al. (2007), and Mitton and Vorkink (2007) imply lower average returns for stocks with higher idiosyncratic skewness. In Barberis and Huang (2008), cumulative prospect theory investors' preference for positively skewed assets results in these assets earning lower average returns. In Brunnermeier et al. (2007), investors optimally trade-off the ex ante utility derived from, and the ex post pain caused by, their subjective beliefs. As a result, good states with small probabilities—states in which positively skewed assets pay off—earn relatively lower expected returns. In Mitton and Vorkink (2007), investors with a preference for skewness optimally choose to remain underdiversified to

Table 7
Factor analysis of *PSS* and size factors.
We run time-series regressions of the *PSS* and size factors on different factor models. *PSS* denotes the predicted systematic skewness factor, *SMB* is the Fama–French size factor used in the four-factor model, and SMB_{FF5} is the Fama–French size factor used in the five-factor model. We use the CAPM in Panel A, the four-factor model with *MKT, SMB*, value (*HML*), and momentum (*MOM*) factors in Panel B, and the five-factor model with *MKT*, SMB_{FF5}, *HML*, profitability (*RMW*), and investment (*CMA*) factors in Panel C. For each regression, we report the monthly α in %, the factor exposures, and adjusted R^2. We report in parentheses the *t*-ratios using a Newey–West estimator with $T^{0.25} \approx 6$ lags. The data are monthly from July 1963 to December 2017.

Factor	α(%)	β_{MKT}	β_{PSS}	β_{SMB}	β_{HML}	β_{MOM}	β_{RMW}	β_{CMA}	Adj. R^2
Panel A: CAPM									
PSS	0.33	0.21							0.05
	(2.22)	(3.98)							
SMB	0.11	0.20							0.08
	(0.90)	(6.71)							
SMB_{FF5}	0.15	0.19							0.07
	(1.24)	(5.73)							
Panel B: Four-factor model									
PSS	−0.03	0.07		0.96	0.06	0.33			0.64
	(−0.27)	(1.70)		(13.42)	(0.93)	(6.02)			
PSS	0.12	0.24			−0.07	0.33			0.17
	(0.86)	(5.15)			(−0.49)	(3.46)			
SMB	0.16	0.18			−0.13	0.01			0.10
	(1.32)	(5.57)			(−1.35)	(0.12)			
SMB	0.09	0.04	0.59		−0.09	−0.19			0.61
	(0.96)	(1.60)	(14.30)		(−2.15)	(−5.16)			
SMB_{FF5}	0.14	0.19			0.00	0.01			0.07
	(1.21)	(5.66)			(0.05)	(0.10)			
SMB_{FF5}	0.07	0.05	0.58		0.04	−0.19			0.57
	(0.81)	(1.98)	(13.65)		(0.91)	(−5.23)			
Panel C: Five-factor model									
PSS	0.49	−0.06		0.83	−0.05		−0.49	−0.28	0.57
	(3.38)	(−1.92)		(12.29)	(−0.59)		(−3.22)	(−1.66)	
PSS	0.75	0.05			−0.01		−0.84	−0.38	0.25
	(4.27)	(0.86)			(−0.04)		(−3.24)	(−1.58)	
SMB	0.33	0.12			−0.09		−0.50	−0.09	0.21
	(2.52)	(2.79)			(−0.82)		(−3.40)	(−0.66)	
SMB	−0.04	0.09	0.51		−0.09		−0.07	0.10	0.55
	(-0.40)	(3.80)	(11.94)		(−1.52)		(−1.21)	(1.16)	
SMB_{FF5}	0.30	0.13			0.06		−0.43	−0.12	0.16
	(2.30)	(3.07)			(0.51)		(−3.03)	(−0.83)	
SMB_{FF5}	−0.08	0.10	0.51		0.06		0.01	0.08	0.52
	(−0.75)	(4.35)	(11.37)		(0.98)		(0.12)	(0.88)	

preserve their portfolio skewness. As a result, positively skewed assets earn lower expected returns. We empirically investigate in this section whether these effects are important in a large cross-section of stock returns.

4.1. What predicts idiosyncratic skewness?

First, we run predictive panel regressions as in Eq. (4) but replacing coskewness ranks on the left-hand side by the cross-sectional ranks of idiosyncratic skewness. We measure stock *i*'s idiosyncratic skewness by estimating the three-moment CAPM using daily excess returns over 12-month periods and computing the skewness of residuals. See Appendix A for more details.

As shown in Bali et al. (2016), the idiosyncratic skewness of individual stocks' Fama and French (1993) three-factor model residuals are very similar to their total skewness, suggesting that the exposures to the market, size, and value factors do not capture the systematic part of skewness. Hence, analyzing the effect of the idiosyncratic skewness of three-factor model residuals is tantamount to analyzing total skewness, which includes systematic skewness. In this section, we instead regress daily returns on the market portfolio returns and its squared values to remove systematic skewness from returns. As a robustness check, we repeat our analysis for total skewness instead of idiosyncratic skewness in Section 4.3.1.

We report in Fig. 3 the time series of regression coefficients and in Table 2 of the Online Appendix their time-series averages and 5th and 95th percentiles obtained by estimating panel regression (4) each month. Fig. 3 can be directly compared to Fig. 2, and Table 2 has the same structure as Table 1 in the Online Appendix.

The rank of idiosyncratic skewness seems more persistent than for coskewness; its coefficient is always positive and, on average, 0.135 (compared to 0.059 for coskewness). Idiosyncratic volatility is also a strong predictor of idiosyncratic skewness as shown in BMV, although its coefficient is briefly negative at the start of the sample period. In contrast to BMV, however, we find that the idiosyncratic skewness rank is a much stronger predictor than idiosyncratic volatility rank, with an average coefficient twice as large.

There is a strong and negative relation between firm size and idiosyncratic skewness, in line with BMV, CHS,

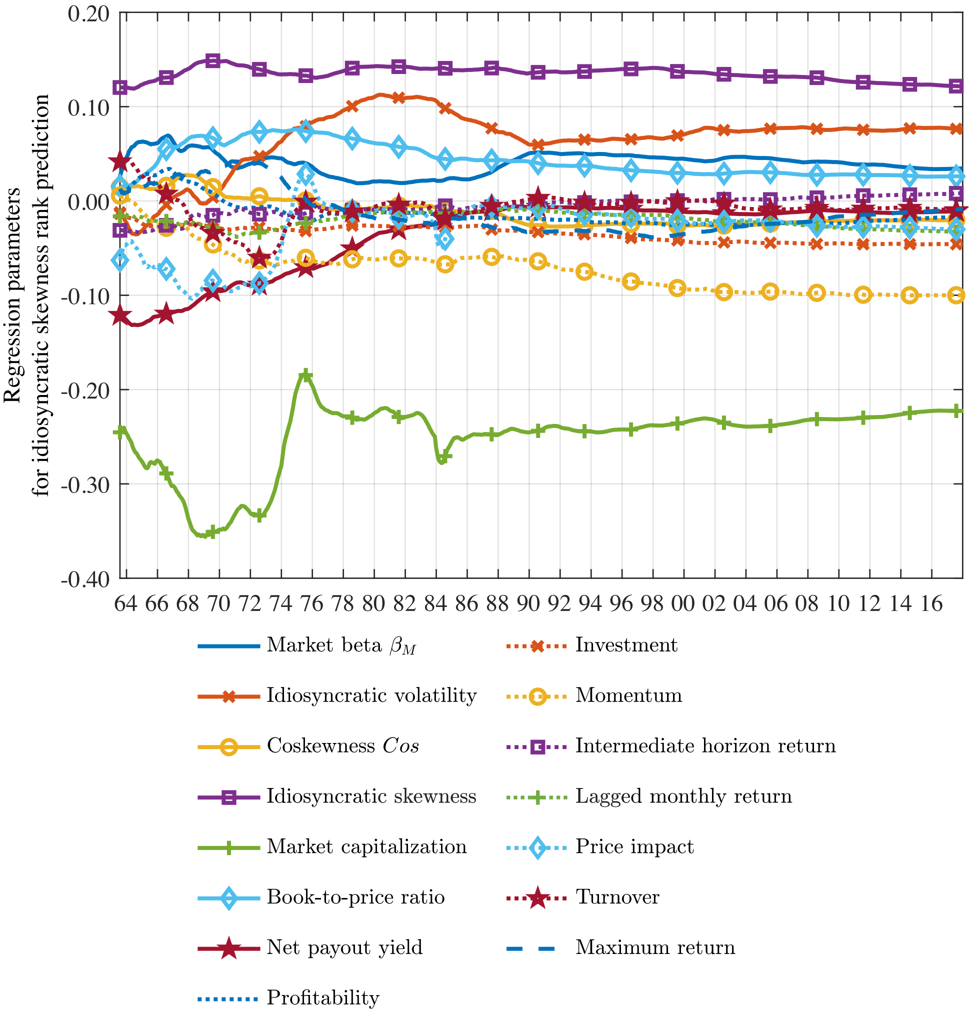

Fig. 3. Coefficients of predictive panel regressions for idiosyncratic skewness ranks. We report the panel regression coefficients $\hat{\theta}$ and $\hat{\phi}$ in Eq. (4) from July 1963 to December 2017. Each month, we run a panel regression that predicts the next 12-month realized daily idiosyncratic skewness using past risk measures and stock characteristics. We use the normalized cross-sectional rank of idiosyncratic skewness as the dependent variable and the normalized cross-sectional ranks of past risk measures and characteristics as predictors. We use all past observations to estimate at each point in time the panel regression. The construction of all variables is detailed in Appendix A.

and Conrad et al. (2013). This is in sharp contrast with coskewness that is positively associated with firm size. Hence, large firms have more positive systematic skewness but more negative idiosyncratic skewness. On the other hand, small firms have a higher likelihood of experiencing large and positive returns, but they also tend to carry more systematic downside risk.

CHS shows that skewness is strongly negatively related to returns over the last six months. We find that coskewness and idiosyncratic skewness are both strongly negatively related to momentum and the lagged monthly return (hence, the return over the last 12 months). The coefficient for intermediate horizon return (from month $t-12$ to $t-7$) is always positive for coskewness, suggesting that returns from month $t-6$ to $t-1$ are more important in predicting coskewness. We do not find the same result for idiosyncratic skewness; the coefficient for intermediate horizon return is close to zero.

Future idiosyncratic skewness ranks are also related to other risk measures and firm characteristics. High β_M and high book-to-price ratio always positively predict idiosyncratic skewness ranks, whereas higher net payout yield and investment predict lower idiosyncratic skewness.

How well do we capture future realized idiosyncratic skewness? In Table 1 we analyzed the realized coskewness of different long-short portfolios and showed that our *PSS* factor generated the most negative realized coskewness. The economics of idiosyncratic skewness is different; it is a potentially priced characteristic, not a risk factor. Skewness-seeking investors modeled in Barberis and Huang (2008), Brunnermeier et al. (2007), and Mitton and Vorkink (2007) are not interested by the idiosyncratic skewness of a portfolio but rather in the high return potential of a single stock. Therefore, we examine the equal-weighted average stock-specific idiosyncratic skewness. This measure conveys the idiosyncratic skewness one can expect by picking one stock among a group of stocks, not the skewness of an equal-weighted diversified portfolio.

Fig. 4 presents evidence on the ability of different predictors in forecasting future realized idiosyncratic skewness. Each month t, we use the predicted cross-sectional idiosyncratic skewness rank to sort stocks into a bottom 30% group and a top 30% group. The thick blue line is the equal-weighted average of realized idiosyncratic skewness from month t to $t+11$ of individual stocks in the bottom group minus the average for stocks in the top group. We also report the average spread using two other idiosyncratic skewness predictors to sort stocks: lagged idiosyncratic volatility and idiosyncratic skewness.

All measures create a negative spread, indicating that the bottom 30% of stocks indeed have lower realized idiosyncratic skewness than the top 30%. The only exception is the November 1986 to August 1987 period during which the spread becomes slightly positive. But most importantly, the predicted idiosyncratic skewness measure creates the lowest (i.e., most negative) difference between the bottom stocks' idiosyncratic skewness and the top stocks. Therefore, our predictive panel regression for cross-sectional ranks is successful in predicting differences across stocks in both realized coskewness and idiosyncratic skewness.[28]

4.2. Factor analysis for idiosyncratic skewness sorted portfolios

In this section, we examine the risk-adjusted performance of idiosyncratic skewness sorted portfolios. We run time-series regressions of idiosyncratic skewness sorted portfolios on the different factor models examined in Section 3. In each panel in Tables 8 and 9, we run a time-series regression for the portfolio that holds each month the bottom 30% of stocks with the lowest predicted idiosyncratic skewness, the portfolio with the middle 40% of stocks, the portfolio with the top 30% of stocks, and finally the long-short portfolio that buys the bottom 30% portfolio and short-sells the top 30% portfolio. Table 8 reports on equal-weighted portfolios and Table 9 on value-weighted portfolios. We use the best model in each of Tables 2–4: Panel A reports on the CAPM with *MKT* augmented with *PSS*; Panel B on the model with *MKT, PSS, HML*, and *MOM*; and Panel C on the model with *MKT, PSS, HML, RMW*, and *CMA*.

We find that across all three factor models and portfolio weighting, the low-minus-high portfolio has negative and significant loadings on the *MKT* and *PSS* factors that vary from −0.39 to −0.14 for *MKT* and from −0.90 to −0.44 for *PSS*. The loadings on *PSS* increase going from the low to the high idiosyncratic skewness portfolios. This is in line with the negative correlation reported in Bali et al. (2016) (see Table 14.5) between coskewness and idiosyncratic skewness. In both Tables, the low-minus-high portfolio has a significantly positive loading on *MOM* in Panel B and on *RMW* in Panel C. Loadings on the *HML* factor are also significantly negative in Panel C.

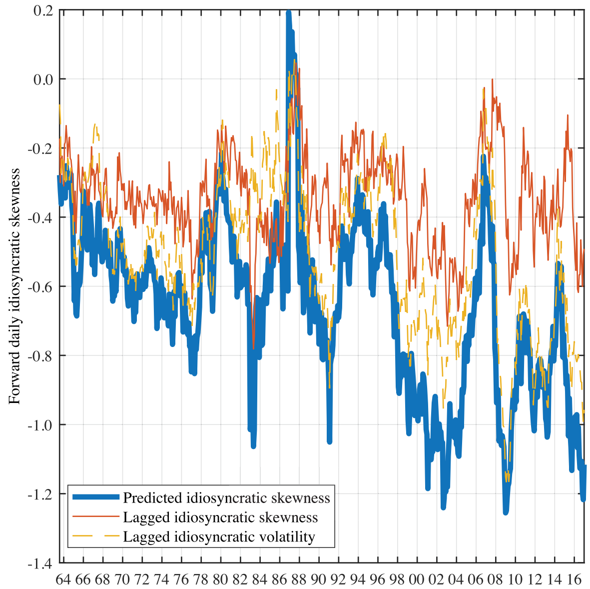

Fig. 4. Equal-weighted average of stock-specific realized idiosyncratic skewness. We report equal-weighted averages of stock-specific realized idiosyncratic skewness. Each month, we rank stocks based on a predictor of future idiosyncratic skewness. As predictors, we use daily return idiosyncratic skewness computed over the last year, daily return idiosyncratic volatility computed over the last year, and the panel regression predicted idiosyncratic skewness cross-sectional ranks. We then compute each stock's daily return idiosyncratic skewness over the next year. For each predictor and each month, we report the equal-weighted average idiosyncratic skewness of the bottom 30% stocks minus the equal-weighted average idiosyncratic skewness of the top 30% stocks. The sample period is July 1963 to December 2017.

If idiosyncratic skewness is priced, then these time-series regressions should reveal a higher α for the low portfolio than for the high portfolio and a positive α for the low-minus-high portfolio.

Across factor models, the regression αs for the equal-weighted low-minus-high idiosyncratic skewness portfolio range from −0.25% to 0.16% per month and are all insignificant. The model with *MKT* and *PSS* in Panel A is not sufficient in all cases: both the equal- and value-weighted Low portfolios and the value-weighted Low-High portfolio have significant αs. However, when we move to other factor models in Panels B and C, all αs become insignificant.

Overall, we show that we can forecast the relative ranking of each stock's future idiosyncratic skewness. The *PSS* is an important factor in explaining the return of idiosyncratic skewness sorted portfolios, but it does not suffice in explaining the higher performance of equal-weighted portfolios with low idiosyncratic skewness and of value-weighted portfolios. But these αs are not robust to the inclusion of other risk factors.

4.3. Robustness checks

In this section, we run robustness checks for the risk-adjusted performance of low idiosyncratic skewness stocks. We examine portfolios sorted by total, instead of idiosyncratic, skewness and sorted by a quantile-based measure of skewness.

[28] The ordering in realized idiosyncratic skewness is robust to using value-weighted averages.

Table 8
Factor analysis of equal-weighted portfolios sorted by predicted idiosyncratic skewness.
We run time-series regressions of portfolio excess returns on different factor models. Each month, we run a panel regression that predicts the next 12-month realized daily idiosyncratic skewness using past risk measures and stock characteristics. We use the cross-sectional rank of idiosyncratic skewness as the dependent variable and the cross-sectional ranks of past risk measures and characteristics as predictors. We form equal-weighted portfolios: one with the bottom 30% stocks with the lowest predicted idiosyncratic skewness ranks (Low), one with the middle 40% stocks (Medium), one with the top 30% stocks with the highest predicted idiosyncratic skewness ranks (High), and a low-minus-high portfolio. We use the CAPM augmented with the predicted systematic skewness factor (*PSS*) in Panel A; the modified four-factor model with *MKT, PSS*, value (*HML*), and momentum (*MOM*) factors in Panel B; and the modified five-factor model with *MKT, PSS, HML*, profitability (*RMW*), and investment (*CMA*) factors in Panel C. For each regression, we report the monthly α in %, the factor exposures, and adjusted R^2. We report in parentheses the *t*-ratios using a Newey–West estimator with $T^{0.25} \approx 6$ lags. The data are monthly from July 1963 to December 2017.

Portfolio	$\alpha(\%)$	β_{MKT}	β_{PSS}	β_{HML}	β_{MOM}	β_{RMW}	β_{CMA}	Adj. R^2
Panel A: MKT, PSS								
Low	0.15	0.91	−0.04					0.92
	(2.43)	(45.06)	(−1.09)					
Medium	0.21	1.04	0.21					0.86
	(2.17)	(36.15)	(3.59)					
High	−0.01	1.19	0.54					0.70
	(−0.04)	(24.83)	(6.88)					
Low-High	0.16	−0.28	−0.59					0.33
	(1.11)	(−5.21)	(−9.46)					
Panel B: MKT, PSS, HML, MOM								
Low	0.05	0.94	−0.04	0.18	0.04			0.93
	(0.99)	(73.09)	(−1.51)	(3.97)	(1.57)			
Medium	0.22	1.03	0.31	0.25	−0.21			0.91
	(2.78)	(48.68)	(8.40)	(6.03)	(−8.34)			
High	0.29	1.09	0.78	0.12	−0.59			0.81
	(1.78)	(23.60)	(13.61)	(1.47)	(−7.53)			
Low-High	−0.25	−0.14	−0.82	0.06	0.63			0.55
	(−1.50)	(−3.13)	(−15.70)	(0.59)	(6.78)			
Panel C: MKT, PSS, HML, RMW, CMA								
Low	−0.06	0.97	0.03	0.09		0.22	0.19	0.94
	(−1.31)	(78.13)	(1.10)	(3.32)		(6.25)	(4.09)	
Medium	0.04	1.10	0.24	0.27		0.02	0.10	0.89
	(0.46)	(52.87)	(5.84)	(5.64)		(0.40)	(1.51)	
High	0.05	1.19	0.46	0.29		−0.44	−0.05	0.73
	(0.23)	(24.66)	(4.65)	(2.90)		(−2.98)	(−0.25)	
Low-High	−0.11	−0.22	−0.44	−0.20		0.67	0.25	0.39
	(−0.46)	(−4.40)	(−4.28)	(−1.89)		(3.96)	(0.00)	

4.3.1. Portfolios sorted by total predicted skewness

In the previous section, we used the skewness of daily residuals from the three-moment CAPM to measure idiosyncratic skewness. To verify that our results are robust to this choice of risk adjustment, we run predictive panel regressions for total skewness. If the *PSS* factor captures systematic skewness risk, then the α of total skewness sorted portfolios, if any, can be attributed to idiosyncratic skewness. On the other hand, if only systematic skewness is priced, then controlling for the exposure to the *PSS* factor should leave no pricing relation for total skewness.

We report in Section 4 of the Online Appendix a figure with time-varying panel regression coefficients, a table with coefficient averages and percentiles, and a figure for the stock-specific realized skewness average of low predicted skewness stocks minus the average for high predicted skewness stocks. These figures and table are directly comparable to Figs. 3 and 4 and to Table 2 in the Online Appendix. Overall, we find that predictors of total skewness are very similar to the ones for idiosyncratic skewness, and the spread in realized total skewness is the lowest across predictors. These results show that our predictive panel regression model is also successful in predicting the cross-sectional ranks in total skewness.

Tables 10 and 11 have the same structure as Tables 8–9 but run time-series regressions for predicted skewness sorted equal-weighted and value-weighted portfolios, respectively.

In line with CHS, who do not find much difference in their predictive model when using either total skewness or market adjusted-return skewness, our results are largely unchanged. The low-minus-high total skewness portfolio has a significantly negative loading on *PSS*. Its α ranges from −0.25% to 0.33% per month across models, which is smaller than the Fama–French three-factor α of slightly more than 1% for low-minus-high risk-neutral skewness portfolios reported in Conrad et al. (2013) (see their Table 4). All low-minus-high portfolio αs are insignificant, except in Panel A for value-weighted portfolios. Therefore, value-weighted skewness portfolios are strongly exposed to *MKT* and *PSS*, but these two factors are not sufficient to explain the higher performance of stocks with low idiosyncratic skewness. When controlling for other risk factors, the risk-adjusted performance of low skewness stocks is not distinguishable from zero.

Table 9
Factor analysis of value-weighted portfolios sorted by predicted idiosyncratic skewness.
We run time-series regressions of portfolio excess returns on different factor models. Each month, we run a panel regression that predicts the next 12-month realized daily idiosyncratic skewness using past risk measures and stock characteristics. We use the cross-sectional rank of idiosyncratic skewness as the dependent variable and the cross-sectional ranks of past risk measures and characteristics as predictors. We form value-weighted portfolios: one with the bottom 30% stocks with the lowest predicted idiosyncratic skewness ranks (Low), one with the middle 40% stocks (Medium), one with the top 30% stocks with the highest predicted idiosyncratic skewness ranks (High), and a low-minus-high portfolio. We use the CAPM augmented with the predicted systematic skewness factor (*PSS*) in Panel A; the modified four-factor model with *MKT, PSS*, value (*HML*), and momentum (*MOM*) factors in Panel B; and the modified five-factor model with *MKT, HML, PSS*, profitability (*RMW*), and investment (*CMA*) factors in Panel C. For each regression, we report the monthly α in %, the factor exposures, and adjusted R^2. We report in parentheses the t-ratios using a Newey–West estimator with $T^{0.25} \approx 6$ lags. The data are monthly from July 1963 to December 2017.

Portfolio	α(%)	β_{MKT}	β_{PSS}	β_{HML}	β_{MOM}	β_{RMW}	β_{CMA}	Adj. R^2
Panel A: MKT, PSS								
Low	0.08	0.91	−0.14					0.97
	(2.78)	(78.43)	(−12.32)					
Medium	0.04	1.13	0.04					0.86
	(0.43)	(30.63)	(0.75)					
High	−0.21	1.31	0.47					0.75
	(−1.44)	(24.22)	(6.11)					
Low-High	0.28	−0.39	−0.62					0.40
	(1.85)	(−6.36)	(−7.63)					
Panel B: MKT, PSS, HML, MOM								
Low	0.00	0.94	−0.17	0.04	0.09			0.98
	(0.04)	(148.90)	(−16.97)	(2.36)	(8.80)			
Medium	0.07	1.11	0.17	0.29	−0.28			0.93
	(1.08)	(52.78)	(6.13)	(7.96)	(−12.14)			
High	0.05	1.22	0.72	0.24	−0.61			0.87
	(0.36)	(31.05)	(14.01)	(4.13)	(−9.65)			
Low-High	−0.05	−0.28	−0.90	−0.20	0.70			0.69
	(−0.33)	(−6.70)	(−15.95)	(−2.79)	(10.16)			
Panel C: MKT, PSS, HML, RMW, CMA								
Low	−0.02	0.94	−0.10	−0.03		0.15	0.12	0.98
	(−0.50)	(118.45)	(−8.47)	(−1.77)		(6.66)	(3.97)	
Medium	−0.06	1.16	0.03	0.39		−0.16	−0.06	0.90
	(−0.71)	(44.58)	(0.92)	(6.03)		(−3.12)	(−0.81)	
High	−0.18	1.32	0.40	0.45		−0.47	−0.12	0.78
	(−0.93)	(27.57)	(4.60)	(4.66)		(−3.67)	(−0.68)	
Low-High	0.17	−0.38	−0.50	−0.49		0.62	0.24	0.49
	(0.76)	(−7.14)	(−5.21)	(−4.38)		(4.21)	(1.18)	

4.3.2. Predicting quantile-based total and idiosyncratic skewness

In this section, we use a different measure for idiosyncratic skewness and total skewness. Given that the sample skewness estimator is sensitive to large values, we follow Ghysels et al. (2016) and use a quantile-based measure of skewness,

$$QSK(x_t) = \frac{(q_{0.95}(x_t) - q_{0.50}(x_t)) - (q_{0.50}(x_t) - q_{0.05}(x_t))}{q_{0.95}(x_t) - q_{0.05}(x_t)}, \tag{10}$$

where x_t are either daily time-series residuals from the three-moment CAPM to capture idiosyncratic skewness or daily returns to capture total skewness, and $q_{0.05}(x_t)$, $q_{0.50}(x_t)$, and $q_{0.95}(x_t)$ are, respectively, the 5th, 50th, and 95th empirical percentiles of x_t. *QSK* measures the standardized difference between the distance between a top percentile and the median and the distance between the median and a bottom percentile. *QSK* is zero for a symmetric distribution and negative (positive) for a negatively (positively) skewed distribution. The advantage of the quantile-based skewness measure *QSK* is that it is robust to the presence of outliers.

We run predictive panel regressions (4) to predict future quantile-based idiosyncratic and total skewness. We then run time-series regressions for predicted QSK rank-sorted portfolios. Tables 4–7 of the Online Appendix report on equal- and value-weighted portfolios sorted by either quantile-based idiosyncratic skewness or quantile-based total skewness.

The low-minus-high skewness portfolios' loadings on the *PSS* factor range from −0.87 to −0.39 and are highly significant in all cases. As before, loadings on *MOM* and *RMW* are positive and significant everywhere.

The αs when controlling for *MKT* and *PSS* are positive and significant, ranging from 0.37% to 0.58% per month. Similar to the results using predicted idiosyncratic or total skewness, *PSS* does not capture all of the outperformance. Low-High portfolio αs are also significant for value-weighted portfolios in Panel B of Table 5 and 7. In contrast, none of the low-minus-high portfolio αs are significantly positive in Panel C of Tables 4 to 7 of the Online Appendix. Therefore, we find stronger evidence

Table 10
Factor analysis of equal-weighted portfolios sorted by predicted skewness.
We run time-series regressions of portfolio excess returns on different factor models. Each month, we run a panel regression that predicts the next 12-month realized daily skewness using past risk measures and stock characteristics. We use the cross-sectional rank of skewness as the dependent variable and the cross-sectional ranks of past risk measures and characteristics as predictors. We form equal-weighted portfolios: one with the bottom 30% stocks with the lowest predicted skewness ranks (Low), one with the middle 40% stocks (Medium), one with the top 30% stocks with the highest predicted skewness ranks (High), and a low-minus-high portfolio. We use the CAPM augmented with the predicted systematic skewness factor (*PSS*) in Panel A; the modified four-factor model with *MKT, PSS*, value (*HML*), and momentum (*MOM*) factors in Panel B; and the modified five-factor model with *MKT, PSS, HML*, profitability (*RMW*), and investment (*CMA*) factors in Panel C. For each regression, we report the monthly α in %, the factor exposures, and adjusted R^2. We report in parentheses the *t*-ratios using a Newey–West estimator with $T^{0.25} \approx 6$ lags. The data are monthly from July 1963 to December 2017.

Portfolio	$\alpha(\%)$	β_{MKT}	β_{PSS}	β_{HML}	β_{MOM}	β_{RMW}	β_{CMA}	Adj. R^2
Panel A: MKT, PSS								
Low	0.16	0.91	−0.04					0.92
	(2.63)	(42.80)	(−0.94)					
Medium	0.22	1.04	0.21					0.86
	(2.15)	(35.77)	(3.80)					
High	−0.02	1.19	0.53					0.69
	(−0.12)	(24.44)	(6.59)					
Low-High	0.18	−0.28	−0.57					0.31
	(1.25)	(−4.96)	(−9.03)					
Panel B: MKT, PSS, HML, MOM								
Low	0.04	0.95	−0.04	0.20	0.05			0.93
	(0.83)	(73.57)	(−1.52)	(4.41)	(2.37)			
Medium	0.23	1.04	0.31	0.25	−0.21			0.91
	(2.83)	(48.49)	(8.74)	(5.91)	(−8.67)			
High	0.29	1.08	0.76	0.11	−0.60			0.80
	(1.76)	(23.04)	(13.21)	(1.37)	(−7.70)			
Low-High	−0.25	−0.13	−0.81	0.09	0.66			0.55
	(−1.54)	(−2.79)	(−15.45)	(0.81)	(7.16)			
Panel C: MKT, PSS, HML, RMW, CMA								
Low	−0.06	0.97	0.04	0.11		0.24	0.19	0.94
	(−1.34)	(78.43)	(1.51)	(3.86)		(6.70)	(3.97)	
Medium	0.05	1.10	0.25	0.26		0.02	0.11	0.88
	(0.50)	(51.79)	(5.97)	(5.40)		(0.37)	(1.53)	
High	0.05	1.19	0.45	0.29		−0.45	−0.05	0.72
	(0.20)	(24.13)	(4.44)	(2.80)		(−3.05)	(−0.25)	
Low-High	−0.10	−0.22	−0.41	−0.18		0.69	0.24	0.38
	(−0.43)	(−4.15)	(−3.98)	(−1.67)		(4.08)	(0.97)	

of low skewness portfolios outperforming high skewness portfolios when predicting a quantile-based measure of skewness, but this outperformance is not robust to the inclusion other risk factors.

5. Monte Carlo evidence

In this section, we provide a Monte Carlo analysis of the added value of our methodology. We simulate data, make predictions using either the raw values of variables or the normalized cross-sectional ranks in a predictive panel regression, and compare their rank correlation with the realized values of the predicted variable. We simulate a sample with $T = 300$ time periods that is approximately half the estimation sample length used in the empirical section. We simulate $N = 2000$ stocks and $K = 10$ predictors.

For each Monte Carlo index $m = 1, \dots, 10,000$, we simulate moments of variables X and Y in Step 1 as described below, values for each variable in Step 2, and compute predictions using either raw values or normalized ranks in Step 3. More precisely,

1. We simulate moments of the variables Xs and Y by
   (a) simulating a correlation matrix by generating a $(K+1) \times (K+1)$ matrix $\tilde{C}$ of random Gaussian variables, normalizing each row such that they have unit length $\bar{C}_{i,j} = \frac{\tilde{C}_{i,j}}{\sqrt{\sum_{l=1}^{K+1} \tilde{C}_{i,l}^2}}$, and setting $C = \bar{C}\bar{C}^\top$ (see Marsaglia and Olkin, 1984);
   (b) simulating expected values of Y_t and each $X_{t,k}$, $t = 1, \dots, T+1$, $k = 1, \dots, K$, as random Gaussian variables;
   (c) and simulating variances of Y_t and each $X_{t,k}$ as chi-square random variables with one degree of freedom.
2. Then, we simulate a sample of values for X_t and Y_{t+1} as coming from the joint normal distribution for $t = 1, \dots, T$ using the correlation matrix, expected values, and variances from Step 1.
3. Finally, we compute two sets of predictions:
   (a) the first set as $X_T \hat{\beta}$ where $\hat{\beta} = (X'X)^{-1}X'Y$, $X = (X_1', \dots, X_{T-1}')'$, and $Y = (Y_2', \dots, Y_T')'$;
   (b) the second set as $U_{X_T} \hat{\beta}_F$ using normalized ranks $U_{Y_t} = F(Y_t)$ and $U_{X_{t,k}} = F(X_{t,k})$ where

Table 11
Factor analysis of value-weighted portfolios sorted by predicted skewness.
We run time-series regressions of portfolio excess returns on different factor models. Each month, we run a panel regression that predicts the next 12-month realized daily skewness using past risk measures and stock characteristics. We use the cross-sectional rank of skewness as the dependent variable and the cross-sectional ranks of past risk measures and characteristics as predictors. We form value-weighted portfolios: one with the bottom 30% stocks with the lowest predicted skewness ranks (Low), one with the middle 40% stocks (Medium), one with the top 30% stocks with the highest predicted skewness ranks (High), and a low-minus-high portfolio. We use the CAPM augmented with the predicted systematic skewness factor (*PSS*) in Panel A; the modified four-factor model with *MKT, PSS*, value (*HML*), and momentum (*MOM*) factors in Panel B; and the modified five-factor model with *MKT, PSS, HML*, profitability (*RMW*), and investment (*CMA*) factors in Panel C. For each regression, we report the monthly α in %, the factor exposures, and adjusted R^2. We report in parentheses the t-ratios using a Newey–West estimator with $T^{0.25} \approx 6$ lags. The data are monthly from July 1963 to December 2017.

Portfolio	$\alpha(\%)$	β_{MKT}	β_{PSS}	β_{HML}	β_{MOM}	β_{RMW}	β_{CMA}	Adj. R^2
Panel A: MKT, PSS								
Low	0.08	0.91	−0.14					0.97
	(2.93)	(77.17)	(−12.02)					
Medium	−0.00	1.14	0.07					0.86
	(−0.02)	(29.28)	(1.41)					
High	−0.25	1.31	0.45					0.73
	(−1.65)	(23.13)	(5.74)					
Low-High	0.33	−0.40	−0.59					0.37
	(2.05)	(−6.14)	(−7.29)					
Panel B: MKT, PSS, HML, MOM								
Low	−0.00	0.94	−0.17	0.05	0.09			0.98
	(−0.02)	(149.20)	(−16.87)	(2.92)	(9.07)			
Medium	0.07	1.11	0.21	0.26	−0.32			0.94
	(1.14)	(53.31)	(7.95)	(6.70)	(−13.81)			
High	0.05	1.20	0.71	0.21	−0.65			0.86
	(0.36)	(30.42)	(13.63)	(3.60)	(−9.89)			
Low-High	−0.05	−0.27	−0.88	−0.16	0.75			0.69
	(−0.34)	(−6.40)	(−15.68)	(−2.29)	(10.66)			
Panel C: MKT, PSS, HML, RMW, CMA								
Low	−0.02	0.94	−0.10	−0.02		0.15	0.12	0.98
	(−0.51)	(117.89)	(−8.51)	(−1.32)		(6.91)	(4.03)	
Medium	−0.07	1.17	0.05	0.37		−0.20	−0.06	0.89
	(−0.78)	(40.99)	(1.37)	(5.60)		(−3.58)	(−0.85)	
High	−0.19	1.32	0.36	0.43		−0.52	−0.13	0.77
	(−0.94)	(26.57)	(4.12)	(4.40)		(−4.05)	(−0.72)	
Low-High	0.18	−0.38	−0.46	−0.46		0.67	0.25	0.46
	(0.77)	(−6.85)	(−4.78)	(−4.11)		(4.57)	(1.21)	

$\hat{\beta}_F = (U_X'U_X)^{-1}U_X'U_Y$, $U_X = (U_{X_1}', \ldots, U_{X_{T-1}}')'$, and $U_Y = (U_{Y_2}', \ldots, U_{Y_T}')'$.

(c) Finally, we compute the predictive rank correlations between predicted and realized values, $\rho_m^S(Y_{T+1}, X_T\hat{\beta})$ and $\rho_{F,m}^S(Y_{T+1}, U_{X_T}\hat{\beta}_F)$.

We start with the case in which the X_{t-1} and Y_t are jointly normally distributed. The top graph in Fig. 5 reports for each simulation m the predictive rank correlation using raw values, $\rho_m^S(Y_{T+1}, X_T\hat{\beta})$, on the horizontal axis and the predictive rank correlation using normalized ranks, $\rho_{F,m}^S(Y_{T+1}, U_{X_T}\hat{\beta}_U)$, on the vertical axis. Points above (below) the 45-degree red line mean that we achieve a higher (lower) predictive rank correlation using normalized ranks instead of raw values. All but a few of the 10,000 predictive rank correlations are higher based on normalized ranks instead of raw values.

To examine the impact of non-normal distributions, we repeat the Monte Carlo analysis by simulating X and Y in Step 4 from a multivariate Student t distribution with six degrees of freedom. The expected values, variances, and correlations are simulated as in Step 1 above. We report in the second graph the predictive rank correlations. As in the normally distributed case, all but a few of the correlation pairs are above the 45-degree line.

Finally, we examine the impact of measurement errors. For each time period, we simulate an additional normally distributed random shock with a volatility equal to 10% of the variable's own volatility. We simulate measurement errors for all variables except for the predicted variable, Y_{T+1}, used to compute predictive rank correlations. Therefore, measurement errors affect the estimation of β and β_F but not the variable that we want to predict. The bottom graph in Fig. 5 shows that predictive rank correlations based on normalized ranks are still higher than those based on raw values.

To summarize, we show by simulation that converting predictors and the predictive variable to normalized cross-sectional ranks leads to higher predictive rank correlations.

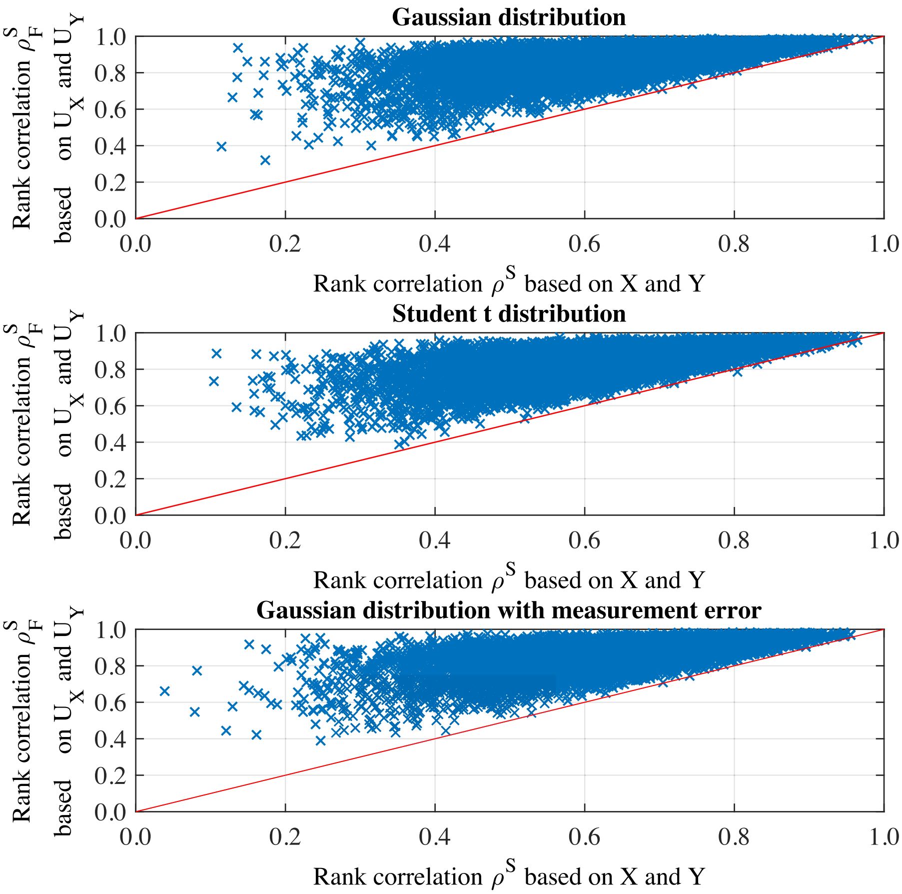

Fig. 5. Monte Carlo simulation—comparison of rank correlations. We report Monte Carlo simulation results on predictive rank correlations. We simulate $m = 10,000$ samples of $T = 300$ time periods with $N = 2000$ stocks and $K = 10$ predictors X. The expected values, variances, and correlation matrix of X and Y are simulated for each sample. Expected values and variances vary over t. The joint distribution of X_t and Y_{t+1} is normal in the top graph, Student t with six degrees of freedom in the middle graph, and normal with measurement errors in the bottom graph. Measurement errors in the bottom graph are simulated for each X and Y using a normal shock with a volatility equal to 10% of the volatility of the respective variable at each point in time. For each sample m, we run a predictive panel linear regression of Y on X and report the predictive rank correlation on the horizontal axis. We also run a predictive panel linear regression using normalized cross-sectional ranks U_Y and U_X and report the predictive rank correlation on the vertical axis. The red line denotes the 45-degree line.

6. Conclusion

We provide a novel empirical methodology to predict future differences in systematic and idiosyncratic skewness across stocks. We form a new systematic skewness risk factor and find that it has a robust and economically sizable risk premium. Finally, we find that idiosyncratic skewness sorted portfolios have a significantly negative loading on idiosyncratic skewness. While the idiosyncratic skewness α is not fully explained by its exposure to the systematic skewness factor, it is not robust to the inclusion of other risk factors such as momentum and profitability.

Our results are important for understanding the relative impact of systematic and idiosyncratic skewness on asset prices. We have relied on models that link risk measures to expected returns and showed which variables best predict these risk measures. A natural extension of our research is to come up with microfoundations for the link between the identified firm characteristics and ex ante risk measures. We leave this aspect for future research.

Appendix A. Data construction

In this Appendix, we detail our data construction. We use market data from the Center for Research in Security Prices (CRSP) that we merge with accounting data from Compustat. We use daily and monthly delisting-adjusted returns for all common stocks with a share code of 10 or 11, and we use the one-month US T-bill rate as the risk-free rate.

For each stock i, we construct the following risk measures. For all measures, we use daily data for days t_d in a 12-month period.

- β_M: We estimate a stock $\beta_{M,i,t\to t+11}$ by running a regression of daily excess returns on a constant and the excess returns on the value-weighted CRSP market portfolio:

$$r_{i,t_d} - r_{f,t_d} = \alpha_{i,t\to t+11} + \beta_{M,i,t\to t+11}\left(r_{M,t_d} - r_{f,t_d}\right) + \epsilon_{i,t_d}. \tag{A.1}$$

- Idiosyncratic volatility: Volatility of the CAPM regression residuals ϵ_{i,t_d} in Eq. (A.1).
- Coskewness *Cos*: We measure coskewness as the covariance between r_{i,t_d} and r^2_{M,t_d}.
- β_{M^2}: We estimate a stock $\beta_{M^2,i,t\to t+11}$ by running a regression of daily excess returns on a constant, the excess returns on the value-weighted CRSP market portfolio, and its square:

$$r_{i,t_d} - r_{f,t_d} = \alpha_{i,t\to t+11} + \beta_{M,i,t\to t+11}\left(r_{M,t_d} - r_{f,t_d}\right) + \beta_{M^2,i,t\to t+11}\left(r_{M,t_d} - r_{f,t_d}\right)^2 + \upsilon_{i,t_d}. \quad (A.2)$$

- β_{HS}: We compute the average of $\epsilon_{i,t_d} \times \epsilon^2_{M,t_d}$ where $\epsilon_{M,t_d} = r_{M,t_d} - \frac{1}{T_{d,t\to t+11}}\sum_{t_d=1}^{T_{d,t\to t+11}} r_{M,t_d}$, and $T_{d,t\to t+11}$ is the number of daily returns from month t to month $t+11$. We divide by the square root of the average of ϵ^2_{i,t_d} times the average of ϵ^2_{M,t_d}.
- Skewness: Mean of the cubed standardized daily returns r_{i,t_d}.
- Quantile-based skewness: We measure robust skewness as in Eq. (10) with daily returns r_{i,t_d}.
- Idiosyncratic skewness: Mean of the cubed standardized regression residuals υ_{i,t_d} in Eq. (A.2).
- Quantile-based idiosyncratic skewness: We measure robust idiosyncratic skewness as in Eq. (10) with daily residuals υ_{i,t_d}.

We construct the following firm characteristics. We impose a six-month lag on all accounting data to ensure that data was available at each point in time.

- Market capitalization: Number of shares outstanding multiplied by the stock price.
- Book-to-price ratio: We measure book-to-price ratio as in Asness and Frazzini (2013). For book value of equity, we use in order of availability stockholder's equity, the sum of common equity and preferred stocks, or total assets minus the sum of total liabilities, minority interest, and preferred stocks. We divide by common shares outstanding or, if it is not available, the sum of shares outstanding for all company issues with an earnings participation flag. We divide the book value per share by the most recent stock price. We set to missing if either the book equity is negative or the stock price is missing.
- Net payout yield: We measure net payout as in Boudoukh et al. (2007). We compute the sum of common stock dividends and the purchase of common and preferred stocks minus the sale of common and preferred stocks. We divide total payout by the most recent market value of equity.
- Profitability-to-asset ratio: We divide gross profitability by total assets.
- Investment: We measure total asset growth on an annual basis.
- Momentum: Total return from month $t-12$ to month $t-2$.
- Intermediate horizon return: Total return from month $t-12$ to month $t-7$.
- Lagged monthly return: Total return for month $t-1$.
- Price impact: The absolute daily returns divided by daily dollar volume averaged over all days in a month for which we have at least five observations.
- Turnover: The sum of dollar volume during a month divided by the market capitalization at the end of the previous month.
- Maximum return: The average of the highest five daily returns in a given month, for months in which there are at least 15 daily returns.

References

Albuquerque, R., 2012. Skewness in stock returns: reconciling the evidence on firm versus aggregate returns. Rev. Financ. Stud. 25, 1630–1673.

Amaya, D., Christoffersen, P., Jacobs, K., Vasquez, A., 2015. Does realized skewness predict the cross-section of equity returns? J. Financ. Econ. 118, 135–167.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. J. Financ. Mark. 5, 31–56.

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. Rev. Financ. Stud. 19, 1191–1239.

Asness, C.S., Frazzini, A., 2013. The devil in HML's details. J. Portf. Manag. 39, 49–68.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. Rev. Financ. Stud. 16, 101–143.

Bali, T.G., Brown, S.J., Murray, S., Tang, Y., 2017. A lottery-demand-based explanation of the beta anomaly. J. Financ. Quant. Anal. 52, 2369–2397.

Bali, T.G., Cakici, N., Whitelaw, R.F., 2011. Maxing out: stocks as lotteries and the cross-section of expected returns. J. Financ. Econ. 99, 427–446.

Bali, T.G., Engle, R.F., Murray, S., 2016. Empirical Asset Pricing. John Wiley and Sons, Hoboken.

Bali, T.G., Murray, S., 2013. Does risk-neutral skewness predict the cross–section of equity option portfolio returns? J. Financ. Quant. Anal. 48, 1145–1171.

Barberis, N., Huang, M., 2008. Stocks as lotteries: the implications of probability weighting for security prices. Am. Econ. Rev. 98, 2066–2100.

Barillas, F., Shanken, J., 2016. Which alpha? Rev. Financ. Stud. 30, 1316–1338.

Boudoukh, J., Michaely, R., Richardson, M., Roberts, M.R., 2007. On the importance of measuring payout yield: implications for empirical asset pricing. J. Financ. 62, 877–915.

Boyer, B., Mitton, T., Vorkink, K., 2010. Expected idiosyncratic skewness. Rev. Financ. Stud. 23, 169–202.

Boyer, B.H., Vorkink, K., 2014. Stock options as lotteries. J. Financ. 69, 1485–1527.

Brunnermeier, M.K., Gollier, C., Parker, J.A., 2007. Optimal beliefs, asset prices, and the preferences for skewed returns. Am. Econ. Rev. 97, 159–165.

Carhart, M.M., 1997. On persistence in mutual fund performance. J. Financ. 52, 57–82.

Chabi-Yo, F., Leisen, D.P., Renault, E., 2014. Aggregation of preferences for skewed asset returns. J. Econ. Theory 154, 453–489.

Chen, J., Hong, H., Stein, J.C., 2001. Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. J. Financ. Econ. 61, 345–381.

Christoffersen, P., Fournier, M., Jacobs, K., Karoui, M., 2017. Option-based estimation of the price of co-skewness and co-kurtosis risk. Unpublished working paper. University of Toronto

Conrad, J., Dittmar, R.F., Ghysels, E., 2013. Ex ante skewness and expected stock returns. J. Financ. 68, 85–124.

Cosemans, M., Frehen, R., Schotman, P.C., Bauer, R., 2015. Estimating security betas using prior information based on firm fundamentals. Rev. Financ. Stud. 29, 1072–1112.

Cremers, M., Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. J. Financ. Quant. Anal. 45, 335–367.

Dahlquist, M., Farago, A., Tedongap, R., 2017. Asymmetries and portfolio choice. Rev. Financ. Stud. 30, 667–702.

DeMiguel, V., Plyakha, Y., Uppal, R., Vilkov, G., 2013. Improving portfolio selection using option-implied volatility and skewness. J. Financ. Quant. Anal. 48, 1813–1845.

Dittmar, R.F., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. J. Financ. 57, 369–403.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.
Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116, 1–22.
Freyberger, J., Neuhierl, A., Weber, M., 2017. Dissecting characteristics non-parametrically. Unpublished working paper. University of Chicago
Ghysels, E., Plazzi, A., Valkanov, R., 2016. Why invest in emerging markets? The role of conditional return asymmetry. J. Financ. 71, 2145–2192.
Giglio, S., Xiu, D., 2017. Inference on risk premia in the presence of omitted factors. Unpublished working paper. University of Chicago
Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21, 1455–1508.
Harvey, C.R., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. Rev. Financ. Stud. 29, 5–68.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J. Financ. 55, 1263–1295.
Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. J. Financ. 45, 881–898.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J. Financ. 48, 65–91.
Kan, R., Robotti, C., Shanken, J., 2013. Pricing model performance and the two-pass cross-sectional regression methodology. J. Financ. 68, 2617–2649.
Kraus, A.K., Litzenberger, R.H., 1976. Skewness preference and the valuation of risk assets. J. Financ. 31, 1085–1100.
Leland, H.E., 1999. Beyond mean-variance: performance measurement in a nonsymmetrical world. Financ. Analyst J. 55, 27–36.
Lettau, M., Maggiori, M., Weber, M., 2014. Conditional risk premia in currency markets and other asset classes. J. Financ. Econ. 114, 197–225.
Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. J. Financ. Econ. 96, 175–194.
Marsaglia, G., Olkin, I., 1984. Generating correlation matrices. SIAM J. Sci. Stat. Comput. 5, 470–475.
Mitton, T., Vorkink, K., 2007. Equilibrium underdiversification and the preference for skewness. Rev. Financ. Stud. 20, 1255–1288.
Novy-Marx, R., 2012. Is momentum really momentum? J. Financ. Econ. 103, 429–453.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. J. Financ. Econ. 108, 1–28.
Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. J. Polit. Econ. 111, 642–685.
Rapach, D., Strauss, J., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Rev. Financ. Stud. 23, 821–862.
Rehman, Z., Vilkov, G., 2012. Risk-neutral skewness: return predictability and its sources. Unpublished working paper. Frankfurt School of Finance and Management
Rubinstein, M., 1973. The fundametal theorem of parameter-preference security valuation. J. Financ. Quant. Anal. 8, 61–69.
Schneider, P., Wagner, C., Zechner, J., 2017. Low risk anomalies?. Unpublished working paper. University of Lugano
Simaan, Y., 1993. Portfolio selection and asset pricing—three-parameter framework. Manag. Sci. 39, 568–577.
Smith, D.R., 2007. Conditional coskewness and asset pricing. J. Empir. Financ. 14, 91–119.
Xing, Y., Zhang, X., Zhao, R., 2010. What does the individual option volatility smirk tell us about future equity returns? J. Financ. Quant. Anal. 45, 641–662.
Zhang, Y., 2005. Individual skewness and the cross-section of average stock returns. Unpublished working paper. Yale University