Journal of Financial Economics 99 (2011) 427–446

Contents lists available at ScienceDirect

Journal of Financial Economics

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

# Maxing out: Stocks as lotteries and the cross-section of expected returns ☆

Turan G. Bali [a,1], Nusret Cakici [b,2], Robert F. Whitelaw [c,d,*]

[a] Department of Economics and Finance, Zicklin School of Business, Baruch College, One Bernard Baruch Way, Box 10-225, New York, NY 10010, United States
[b] Department of Finance, Fordham University, Fordham University, 113 West 60th Street, New York, NY 10023, United States
[c] Stern School of Business, New York University, 44 W. 4th Street, Suite 9-190, New York, NY 10012, United States
[d] NBER, United States

ARTICLE INFO

*Article history:*
Received 24 November 2009
Received in revised form
2 February 2010
Accepted 2 March 2010

*JEL classification:*
G11
G17
G12

*Keywords:*
Extreme returns
Lottery-like payoffs
Cross-sectional return predictability
Skewness preference
Idiosyncratic volatility

ABSTRACT

Motivated by existing evidence of a preference among investors for assets with lottery-like payoffs and that many investors are poorly diversified, we investigate the significance of extreme positive returns in the cross-sectional pricing of stocks. Portfolio-level analyses and firm-level cross-sectional regressions indicate a negative and significant relation between the maximum daily return over the past one month (MAX) and expected stock returns. Average raw and risk-adjusted return differences between stocks in the lowest and highest MAX deciles exceed 1% per month. These results are robust to controls for size, book-to-market, momentum, short-term reversals, liquidity, and skewness. Of particular interest, including MAX reverses the puzzling negative relation between returns and idiosyncratic volatility recently shown in Ang, Hodrick, Xing, and Zhang (2006, 2009).



## 1. Introduction

What determines the cross-section of expected stock returns? This question has been central to modern financial economics since the path breaking work of Sharpe (1964), Lintner (1965), and Mossin (1966). Much of this work has focused on the joint distribution of individual stock returns and the market portfolio as the determinant of expected returns. In the classic capital asset pricing model (CAPM) setting, i.e., with either quadratic preferences or normally distributed returns, expected returns on individual stocks are determined by the covariance of their returns with the market portfolio. Introducing a preference for skewness leads to the three-moment CAPM of Kraus and Litzenberger (1976), which has received empirical support in the literature as, for example, in Harvey and Siddique (2000) and Smith (2007).

Diversification plays a critical role in these models due to the desire of investors to avoid variance risk, i.e., to diversify away idiosyncratic volatility, yet a closer examination of the portfolios of individual investors suggests

☆ We would like to thank Yakov Amihud, Xavier Gabaix, Evgeny Landres, Orly Sade, Jacob Sagi, Daniel Smith, Jeff Wurgler, and seminar participants at the Cesaerea 6th Annual Conference, Arison School of Business, IDC; HEC, Paris; INSEAD; New York University; and Simon Fraser University for helpful comments.

* Corresponding author at: Stern School of Business, New York University, 44 W. 4th Street, Suite 9-190, New York, NY 10012, United States. Tel.: +1 212 998 0338.

*E-mail addresses:* turan.bali@baruch.cuny.edu (T.G. Bali), cakici@fordham.edu (N. Cakici), rwhitela@stern.nyu.edu (R.F. Whitelaw).

[1] Tel.: +1 646 312 3506; fax: +1 646 312 3451.

[2] Tel.: +1 212 636 6776; fax: +1 212 765 5573.

0304-405X/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfineco.2010.08.014

that these investors are, in general, not well-diversified.[3] There may be plausible explanations for this lack of diversification, such as the returns to specialization in information acquisition (Van Nieuwerburgh and Veldkamp, 2010), but nevertheless this empirical phenomenon suggests looking more closely at the distribution of individual stock returns rather than just co-moments as potential determinants of the cross-section of expected returns.

There is also evidence that investors have a preference for lottery-like assets, i.e., assets that have a relatively small probability of a large payoff. Two prominent examples are the favorite-longshot bias in horsetrack betting, i.e., the phenomenon that the expected return per dollar wagered tends to increase monotonically with the probability of the horse winning, and the popularity of lottery games despite the prevalence of negative expected returns (Thaler and Ziemba, 1988). Interestingly, in the latter case, there is increasing evidence that it is the degree of skewness in the payoffs that appeals to participants (Garrett and Sobel, 1999; Walker and Young, 2001), although there are alternative explanations, such as lumpiness in the goods market (Patel and Subrahmanyam, 1978). In the context of the stock market, Kumar (2009) shows that certain groups of individual investors appear to exhibit a preference for lottery-type stocks, which he defines as low-priced stocks with high idiosyncratic volatility and high idiosyncratic skewness.

Motivated by these two literatures, we examine the role of extreme positive returns in the cross-sectional pricing of stocks. Specifically, we sort stocks by their maximum daily return during the previous month and examine the monthly returns on the resulting portfolios over the period July 1962–December 2005. For value-weighted decile portfolios, the difference between returns on the portfolios with the highest and lowest maximum daily returns is $-1.03\%$. The corresponding Fama-French-Carhart four-factor alpha is $-1.18\%$. Both return differences are statistically significant at all standard significance levels. In addition, the results are robust to sorting stocks not only on the single maximum daily return during the month, but also the average of the two, three, four, or five highest daily returns within the month. This evidence suggests that investors may be willing to pay more for stocks that exhibit extreme positive returns, and thus, these stocks exhibit lower returns in the future.

This interpretation is consistent with cumulative prospect theory (Tversky and Kahneman, 1992) as modeled in Barberis and Huang (2008). Errors in the probability weighting of investors cause them to overvalue stocks that have a small probability of a large positive return. It is also consistent with the optimal beliefs framework of Brunnermeier, Gollier, and Parker (2007). In this model, agents optimally choose to distort their beliefs about future probabilities in order to maximize their current utility. Critical to these interpretations of the empirical evidence, stocks with extreme positive returns in a given month should also be more likely to exhibit this phenomenon in the future. We confirm this persistence, showing that stocks in the top decile in one month have a 35% probability of being in the top decile in the subsequent month and an almost 70% probability of being in one of the top three deciles. Moreover, maximum daily returns exhibit substantial persistence in firm-level cross-sectional regressions, even after controlling for a variety of other firm-level variables.

Not surprisingly, the stocks with the most extreme positive returns are not representative of the full universe of equities. For example, they tend to be small, illiquid securities with high returns in the portfolio formation month and low returns over the prior 11 months. To ensure that it is not these characteristics, rather than the extreme returns, that are driving the documented return differences, we perform a battery of bivariate sorts and re-examine the raw return and alpha differences. The results are robust to sorts on size, book-to-market ratio, momentum, short-term reversals, and illiquidity. Results from cross-sectional regressions corroborate this evidence.

Are there alternative interpretations of this apparently robust empirical phenomenon? Recent papers by Ang, Hodrick, Xing, and Zhang (2006, 2009) contain the anomalous finding that stocks with high idiosyncratic volatility have low subsequent returns. It is no surprise that the stocks with extreme positive returns also have high idiosyncratic (and total) volatility when measured over the same time period. This positive correlation is partially by construction, since realized monthly volatility is calculated as the sum of squared daily returns, but even excluding the day with the largest return in the volatility calculation only reduces this association slightly. Could the maximum return simply be proxying for idiosyncratic volatility? We investigate this question using two methodologies, bivariate sorts on extreme returns and idiosyncratic volatility and firm-level cross-sectional regressions. The conclusion is that not only is the effect of extreme positive returns we find robust to controls for idiosyncratic volatility, but that this effect reverses the idiosyncratic volatility effect shown in Ang, Hodrick, Xing, and Zhang (2006, 2009). When sorted first on maximum returns, the equal-weighted return difference between high and low idiosyncratic volatility portfolios is *positive* and both economically and statistically significant. In a cross-sectional regression context, when both variables are included, the coefficient on the maximum return is negative and significant while that on idiosyncratic volatility is *positive*, albeit insignificant in some specifications. These results are consistent with our preferred explanation—poorly diversified investors dislike idiosyncratic volatility, like lottery-like payoffs, and influence prices and hence future returns.

A slightly different interpretation of our evidence is that extreme positive returns proxy for skewness, and investors exhibit a preference for skewness. For example, Mitton and Vorkink (2007) develop a model of agents with heterogeneous skewness preferences and show that

[3] See, for example, Odean (1999), Mitton and Vorkink (2007), and Goetzmann and Kumar (2008) for evidence based on the portfolios of a large sample of U.S. individual investors. Calvet, Campbell, and Sodini (2007) present evidence on the underdiversification of Swedish households, which can also be substantial, although the associated welfare costs for the median household appear to be small.

the result is an equilibrium in which idiosyncratic skewness is priced. However, we show that the extreme return effect is robust to controls for total and idiosyncratic skewness and to the inclusion of a measure of expected skewness as in Boyer, Mitton, and Vorkink (2010). It is also unaffected by controls for co-skewness, i.e., the contribution of an asset to the skewness of a well-diversified portfolio.

The paper is organized as follows. Section 2 provides the univariate portfolio-level analysis, and the bivariate analyses and firm-level cross-sectional regressions that examine a comprehensive list of control variables. Section 3 focuses more specifically on extreme returns and idiosyncratic volatility. Section 4 presents results for skewness and extreme returns. Section 5 concludes.

## 2. Extreme positive returns and the cross-section of expected returns

### 2.1. Data

The first data set includes all New York Stock Exchange (NYSE), American Stock Exchange (Amex), and Nasdaq financial and nonfinancial firms from the Center for Research in Security Prices (CRSP) for the period from January 1926 through December 2005. We use daily stock returns to calculate the maximum daily stock returns for each firm in each month as well as such variables as the market beta, idiosyncratic volatility, and various skewness measures; we use monthly returns to calculate proxies for intermediate-term momentum and short-term reversals; we use volume data to calculate a measure of illiquidity; and we use share prices and shares outstanding to calculate market capitalization. The second data set is Compustat, which is used to obtain the equity book values for calculating the book-to-market ratios of individual firms. These variables are defined in detail in the Appendix and are discussed as they are used in the analysis.

### 2.2. Univariate portfolio-level analysis

Table 1 presents the value-weighted and equal-weighted average monthly returns of decile portfolios that are formed by sorting the NYSE/Amex/Nasdaq stocks based on the maximum daily return within the previous month (MAX). The results are reported for the sample period July 1962–December 2005.

Portfolio 1 (low MAX) is the portfolio of stocks with the lowest maximum daily returns during the past month, and portfolio 10 (high MAX) is the portfolio of stocks with the highest maximum daily returns during the previous month. The value-weighted average raw return difference between decile 10 (high MAX) and decile 1 (low MAX) is $-1.03\%$ per month with a corresponding Newey-West (1987) $t$-statistic of $-2.83$. In addition to the average raw returns, Table 1 also presents the intercepts (Fama-French-Carhart four-factor alphas) from the regression of the value-weighted portfolio returns on a constant, the excess market return, a size factor (SMB), a book-to-market factor (HML), and a momentum factor (MOM), following Fama and French (1993) and Carhart (1997).[4] As shown in the last row of Table 1, the difference in alphas between the high MAX and low MAX portfolios is $-1.18\%$ per month with a Newey–West $t$-statistic of $-4.71$. This difference is economically significant and statistically significant at all conventional levels.

Taking a closer look at the value-weighted average returns and alphas across deciles, it is clear that the pattern is not one of a uniform decline as MAX increases. The average returns of deciles 1–7 are approximately the same, in the range of 1.00–1.16% per month, but, going from decile 7 to decile 10, average returns drop significantly, from 1.00% to 0.86%, 0.52%, and then to $-0.02\%$ per month. The alphas for the first seven deciles are also similar and close to zero, but again they fall dramatically for deciles 8 through 10. Interestingly, the reverse of this pattern is evident across the deciles in the average across months of the average maximum daily return of the stocks within each decile. By definition, this average increases monotonically from deciles 1 to 10, but this increase is far more dramatic for deciles 8, 9, and 10. These deciles contain stocks with average maximum daily returns of 9%, 12%, and 24%, respectively. Given a preference for upside potential, investors may be willing to pay more for, and accept lower expected returns on, assets with these extremely high positive returns. In other words, it is conceivable that investors view these stocks as valuable lottery-like assets, with a small chance of a large gain.

As shown in the third column of Table 1, similar, although somewhat less economically and statistically significant results, are obtained for the returns on equal-weighted portfolios. The average raw return difference between the low MAX and high MAX portfolios is $-0.65\%$ per month with a $t$-statistic of $-1.83$. The corresponding difference in alphas is $-0.66\%$ per month with a $t$-statistic of $-2.31$. As with the value-weighted returns, it is the extreme deciles, in this case deciles 9 and 10, that exhibit low future returns and negative alphas.

For the analysis in Table 1, we start the sample in July 1962 because this starting point corresponds to that used in much of the literature on the cross-section of expected returns; however, the results are similar using the sample starting in January 1926 and for various subsamples. For example, for the January 1926–June 1962 subsample, the average risk-adjusted return difference for the value-weighted portfolios is $-1.25\%$ per month, with a corresponding $t$-statistic of $-3.43$. When we break the original sample at the end of 1983, the subperiods have alpha differences of $-1.62\%$ and $-0.99\%$ per month, both of which are statistically significant. In the remainder of the paper, we continue presenting results for the July 1962–December 2005 sample for comparability with earlier studies.

While conditioning on the single day with the maximum return is both simple and intuitive as a proxy for

[4] Small minus big (SMB), high minus low (HML), and MOM (winner minus loser) are described in and obtained from Kenneth French's data library: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

**Table 1**
Returns and alphas on portfolios of stocks sorted by MAX.

Decile portfolios are formed every month from July 1962 to December 2005 by sorting stocks based on the maximum daily return (MAX) over the past one month. Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) maximum daily returns over the past one month. The table reports the value-weighted (VW) and equal-weighted (EW) average monthly returns, the four-factor Fama-French-Carhart alphas on the value-weighted and equal-weighted portfolios, and the average maximum daily return of stocks within a month. The last two rows present the differences in monthly returns and the differences in alphas with respect to the four-factor Fama-French-Carhart model between portfolios 10 and 1 and the corresponding $t$-statistics. Average raw and risk-adjusted returns, and average daily maximum returns are given in percentage terms. Newey-West (1987) adjusted $t$-statistics are reported in parentheses.

| | VW Portfolios | | EW Portfolios | | |
|---|---|---|---|---|---|
| Decile | Average return | Four-factor alpha | Average return | Four-factor alpha | Average MAX |
| Low MAX | 1.01 | 0.05 | 1.29 | 0.22 | 1.30 |
| 2 | 1.00 | 0.00 | 1.45 | 0.33 | 2.47 |
| 3 | 1.00 | 0.04 | 1.55 | 0.39 | 3.26 |
| 4 | 1.11 | 0.16 | 1.55 | 0.39 | 4.06 |
| 5 | 1.02 | 0.09 | 1.49 | 0.31 | 4.93 |
| 6 | 1.16 | 0.15 | 1.49 | 0.33 | 5.97 |
| 7 | 1.00 | 0.03 | 1.37 | 0.23 | 7.27 |
| 8 | 0.86 | −0.21 | 1.32 | 0.20 | 9.07 |
| 9 | 0.52 | −0.49 | 1.04 | −0.09 | 12.09 |
| High MAX | −0.02 | −1.13 | 0.64 | −0.44 | 23.60 |
| 10-1 | −1.03 | −1.18 | −0.65 | −0.66 | |
| difference | (−2.83) | (−4.71) | (−1.83) | (−2.31) | |

extreme positive returns, it is also slightly arbitrary. As an alternative, we also rank stocks by the average of the $N$ ($N$=1, 2, ..., 5) highest daily returns within the month, with the results reported in Table 2. As before, we report the difference between the returns and alphas on the deciles of firms with the highest and lowest average daily returns over the prior month. For ease of comparison, we report the results from Table 1 in the first column ($N$=1). For both the value-weighted (Panel A) and the equal-weighted portfolios (Panel B), the return patterns when sorting on average returns over multiple days are similar to those when sorting on the single maximum daily return. In fact, if anything, the raw return and alpha differences are both economically and statistically more significant as we average over more days. For example, for value-weighted returns these differences increase in magnitude from −1.03% and −1.18% for $N$=1 to −1.23% and −1.32% for $N$=5.

Another alternative measure of the extent to which a stock exhibits lottery-like payoffs is to compute MAX over longer past periods. Consequently, we first form the MAX(1) portfolios based on the highest daily return over the past 3, 6, and 12 months, and the average raw return differences between the high MAX and low MAX portfolios are −0.63%, −0.52%, and −0.41% per month, respectively. Although these return differences are economically significant, we have statistical significance only for MAX(1) computed over the past quarter. When the MAX(5) portfolios are formed based on the five largest daily returns over the past 3, 6, and 12 months, the average raw return differences are larger (−1.27%, −1.15%, and −0.86% per month, respectively), and they are all statistically significant. More importantly, the differences between the four-factor Fama-French-Carhart alphas for the low and high MAX portfolios are negative and economically and statistically significant for all measures of MAX(1) and MAX(5). Specifically, the alpha differences for the MAX(1) portfolios are in the range of −0.68% to −0.74% per month with $t$-statistics ranging from −2.52 to −2.92. For MAX(5) the results are even stronger, with alpha differences ranging between −1.20% and −1.41% per month and $t$-statistics between −3.78 and −4.36. Finally, we also consider a measure defined as the maximum daily return in a month *averaged* over the past 3, 6, and 12 months. The average raw and risk-adjusted return differences between the extreme portfolios are negative and highly significant without exception.[5]

These analyses show that different proxies for lottery-like payoffs generate similar results, confirming their robustness and thus providing further support for the explanation we offer. For simplicity we focus on MAX(1) over the previous month in the remainder of the paper except in cases where the multiple-day averages are needed to illustrate or illuminate a point.

Of course, the maximum daily returns shown in Table 1 and those underlying the portfolio sorts in Table 2 are for the portfolio formation month, not for the subsequent month over which we measure average returns. Investors may pay high prices for stocks that have exhibited extreme positive returns in the past in the expectation that this behavior will be repeated in the future, but a natural question is whether these expectations are rational. We investigate this issue by examining the average month-to-month portfolio transition matrix, i.e., the average probability that a stock in decile $i$ in one month will be in decile $j$ in the subsequent month (although for brevity, we do not report these results in detail). If maximum daily returns are completely random, then all the probabilities should be approximately 10%,

[5] In the interest of brevity, we do not present detailed results for these alternative measures of MAX, but they are available from the authors upon request.

**Table 2**
Returns on portfolios of stocks sorted by multi-day maximum returns.

Decile portfolios are formed every month from July 1962 to December 2005 by sorting stocks based on the average of the $N$ highest daily returns (MAX($N$)) over the past one month. Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) maximum multi-day returns over the past one month. The table reports the value-weighted (Panel A) and equal-weighted (Panel B) average monthly returns for $N$=1,…,5. The last two rows present the differences in monthly returns and the differences in alphas with respect to the four-factor Fama-French-Carhart model between portfolios 10 and 1. Average raw and risk-adjusted returns are given in percentage terms. Newey-West (1987) adjusted $t$-statistics are reported in parentheses.

| *Panel A: Value-weighted returns on MAX(N) portfolios* | | | | | |
|---|---|---|---|---|---|
| Decile | $N$=1 | $N$=2 | $N$=3 | $N$=4 | $N$=5 |
| Low MAX($N$) | 1.01 | 1.00 | 1.05 | 1.02 | 1.05 |
| 2 | 1.00 | 0.96 | 0.98 | 1.02 | 1.07 |
| 3 | 1.00 | 1.06 | 1.09 | 1.08 | 1.06 |
| 4 | 1.11 | 1.08 | 1.02 | 1.01 | 1.04 |
| 5 | 1.02 | 1.08 | 1.05 | 1.06 | 1.04 |
| 6 | 1.16 | 1.03 | 1.08 | 1.03 | 1.01 |
| 7 | 1.00 | 1.04 | 1.00 | 1.06 | 1.06 |
| 8 | 0.86 | 0.78 | 0.68 | 0.70 | 0.70 |
| 9 | 0.52 | 0.50 | 0.49 | 0.43 | 0.48 |
| High MAX($N$) | −0.02 | −0.16 | −0.13 | −0.12 | −0.18 |
| Return difference | −1.03 | −1.16 | −1.18 | −1.14 | −1.23 |
| | (−2.83) | (−2.97) | (−2.95) | (−2.74) | (−2.93) |
| Alpha difference | −1.18 | −1.29 | −1.26 | −1.21 | −1.32 |
| | (−4.71) | (−4.56) | (−4.12) | (−3.71) | (−4.07) |

| *Panel B: Equal-weighted returns on MAX(N) portfolios* | | | | | |
|---|---|---|---|---|---|
| Decile | $N$=1 | $N$=2 | $N$=3 | $N$=4 | $N$=5 |
| Low MAX($N$) | 1.29 | 1.28 | 1.27 | 1.29 | 1.30 |
| 2 | 1.45 | 1.45 | 1.48 | 1.49 | 1.54 |
| 3 | 1.55 | 1.55 | 1.56 | 1.59 | 1.59 |
| 4 | 1.55 | 1.58 | 1.61 | 1.62 | 1.60 |
| 5 | 1.49 | 1.56 | 1.56 | 1.52 | 1.55 |
| 6 | 1.49 | 1.45 | 1.49 | 1.53 | 1.52 |
| 7 | 1.37 | 1.44 | 1.43 | 1.42 | 1.43 |
| 8 | 1.32 | 1.28 | 1.27 | 1.28 | 1.26 |
| 9 | 1.04 | 1.01 | 1.00 | 0.95 | 0.94 |
| High MAX($N$) | 0.64 | 0.59 | 0.54 | 0.51 | 0.49 |
| Return difference | −0.65 | −0.69 | −0.73 | −0.78 | −0.81 |
| | (−1.83) | (−1.88) | (−1.99) | (−2.11) | (−2.21) |
| Alpha difference | −0.66 | −0.72 | −0.78 | −0.84 | −0.89 |
| | (−2.31) | (−2.36) | (−2.53) | (−2.75) | (−2.93) |

since a high or low maximum return in one month should say nothing about the maximum return in the following month. Instead, there is clear evidence that MAX is persistent, with all the diagonal elements of the transition matrix exceeding 10%. Of greater importance, this persistence is especially strong for the extreme portfolios. Stocks in decile 10 (high MAX) have a 35% chance of appearing in the same decile next month. Moreover, they have a 68% probability of being in deciles 8–10, all of which exhibit high maximum daily returns in the portfolio formation month and low returns in the subsequent month.

A slightly different way to examine the persistence of extreme positive daily returns is to look at firm-level cross-sectional regressions of MAX on lagged predictor variables. Specifically, for each month in the sample we run a regression across firms of the maximum daily return within that month on the maximum daily return from the previous month and seven lagged control variables that are defined in the Appendix and discussed in more detail later—the market beta (BETA), the market capitalization (SIZE), the book-to-market ratio (BM), the return in the previous month (REV), the return over the 11 months prior to that month (MOM), a measure of illiquidity (ILLIQ), and the idiosyncratic volatility (IVOL).[6] Table 3 reports the average cross-sectional coefficients from these regressions and the Newey-West (1987) adjusted $t$-statistics. In the univariate regression of MAX on lagged MAX, the coefficient is positive, quite large, and extremely statistically significant, and the $R$-squared of over 16% indicates substantial cross-sectional explanatory power. In other words, stocks with extreme positive daily returns in one month also tend to exhibit similar features in the following month. When the seven control variables are added to the regression, the coefficient on lagged MAX remains large and significant. Of these seven variables, it

[6] The high cross-sectional correlation between MAX and IVOL, as shown later in Table 9 and discussed in Section 3, generates a multicollinearity problem in the regression; therefore, we orthogonalize IVOL for the purposes of regressions that contain both variables.

**Table 3**
Cross-sectional predictability of MAX.

Each month from July 1962 to December 2005 we run a firm-level cross-sectional regression of the maximum daily return in that month (MAX) on subsets of lagged predictor variables including MAX in the previous month and seven control variables that are defined in the Appendix. The table reports the time-series averages of the cross-sectional regression coefficients, their associated Newey-West (1987) adjusted *t*-statistics (in parentheses), and the regression *R*-squareds.

| MAX | BETA | SIZE | BM | MOM | REV | ILLIQ | IVOL | $R^2$ |
|---|---|---|---|---|---|---|---|---|
| 0.4054 | | | | | | | | 16.64% |
| (45.34) | | | | | | | | |
| | 0.1116 | | | | | | | 1.00% |
| | (4.47) | | | | | | | |
| | | −1.3381 | | | | | | 15.99% |
| | | (−22.42) | | | | | | |
| | | | 0.5334 | | | | | 1.81% |
| | | | (6.41) | | | | | |
| | | | | −1.7264 | | | | 3.52% |
| | | | | (−6.87) | | | | |
| | | | | | −0.0655 | | | 3.31% |
| | | | | | (−11.19) | | | |
| | | | | | | 0.1209 | | 4.28% |
| | | | | | | (8.13) | | |
| | | | | | | | 0.1643 | 27.36% |
| | | | | | | | (86.41) | |
| 0.3325 | 0.2500 | −0.4737 | −0.1277 | −0.3432 | −0.0504 | 0.0200 | 0.1930 | 35.10% |
| (31.31) | (12.14) | (−30.45) | (−5.86) | (−4.47) | (−22.25) | (6.16) | (41.60) | |

is SIZE and IVOL that contribute most to the explanatory power of the regression, with univariate *R*-squareds of 16% and 27%, respectively. The remaining five variables all have univariate *R*-squareds of less than 5%.

As a final check on the return characteristics of stocks with extreme positive returns, we examine more closely the distribution of monthly returns on stocks in the high MAX and low MAX portfolios. Tables 1 and 2 report the mean returns on these stocks, and the cross-sectional regressions in Table 3 and the portfolio transition matrix show that the presence, or absence, of extreme positive returns is persistent, but what are the other features of the return distribution?

Table 4 presents descriptive statistics for the approximately 240,000 monthly returns on stocks within the two extreme deciles in the post-formation month. The mean returns are almost identical to those reported in Table 1 for the equal-weighted portfolio. The slight difference is attributable to the fact that Table 1 reports averages of returns across equal-weighted portfolios that contain slightly different numbers of stocks, whereas Table 4 weights all returns equally. In addition to having a lower average return, high MAX stocks display significantly higher volatility and more positive skewness. The percentiles of the return distribution illustrate the upper tail behavior. While median returns on high MAX stocks are lower, the returns at the 90th, 95th, and 99th percentiles are more than twice as large as those for low MAX stocks. Clearly, high MAX stocks exhibit higher probabilities of extreme positive returns in the following month. The percentiles of the distribution are robust to outliers, but the moments are not, so in the final two columns we report statistics for returns where the 0.5% most extreme returns in both tails have been eliminated. While means, standard deviations, and skewness for the trimmed distributions fall, the relative ordering remains—high MAX stocks have lower means, but higher volatilities and skewness than their low MAX counterparts in the subsequent month.

**Table 4**
Distribution of monthly returns for stocks in the high and low MAX portfolios.

Decile portfolios are formed every month from July 1962 to December 2005 by sorting stocks based on the maximum daily returns (MAX) over the past one month. The table reports descriptive statistics for the approximately 240,000 monthly returns on the individual stocks in deciles 1 (low MAX) and 10 (high MAX) in the following month. The tails of the return distribution are trimmed by removing the 0.5% most extreme observations in each tail prior to the calculation of the statistics in the final two columns.

| | | | Trimmed | |
|---|---|---|---|---|
| | Low MAX | High MAX | Low MAX | High MAX |
| Mean | 1.26% | 0.60% | 1.04% | −0.16% |
| Median | 0.35% | −2.50% | 0.35% | −2.50% |
| Std dev | 12.54% | 30.21% | 9.70% | 24.12% |
| Skewness | 4.26 | 5.80 | 0.59 | 1.35 |
| Percentiles | | | | |
| 1% | −29.6% | −52.1% | | |
| 5% | −14.7% | −33.8% | | |
| 10% | −9.3% | −25.9% | | |
| 25% | −3.4% | −14.3% | | |
| 50% | 0.3% | −2.5% | | |
| 75% | 5.1% | 9.5% | | |
| 90% | 11.6% | 28.6% | | |
| 95% | 17.7% | 46.3% | | |
| 99% | 40.0% | 100.0% | | |

We do not measure investor expectations directly, but the results presented in Tables 3 and 4 are certainly consistent with the underlying theory about preferences for stocks with extreme positive returns. While MAX measures the propensity for a stock to deliver lottery-like

**Table 5**
Summary statistics for decile portfolios of stocks sorted by MAX.

Decile portfolios are formed every month from July 1962 to December 2005 by sorting stocks based on the maximum (MAX) daily returns over the past one month. Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) maximum daily returns over the past one month. The table reports for each decile the average across the months in the sample of the median values within each month of various characteristics for the stocks—the maximum daily return (in percent), the market beta, the market capitalization (in millions of dollars), the book-to-market (BM) ratio, our measure of illiquidity (scaled by $10^5$), the price (in dollars), the return in the portfolio formation month (labeled REV), the cumulative return over the 11 months prior to portfolio formation (labeled MOM), and the idiosyncratic volatility over the past one month (IVOL). There is an average of 309 stocks per portfolio.

| Decile | MAX | Size ($\$10^6$) | Price (\$) | Market beta | BM ratio | Illiquidity ($10^5$) | IVOL | REV | MOM |
|---|---|---|---|---|---|---|---|---|---|
| Low MAX | 1.62 | 316.19 | 25.44 | 0.33 | 0.7259 | 0.2842 | 0.97 | −2.44 | 10.95 |
| 2 | 2.51 | 331.47 | 25.85 | 0.55 | 0.6809 | 0.1418 | 1.26 | −0.96 | 11.16 |
| 3 | 3.22 | 250.98 | 23.88 | 0.68 | 0.6657 | 0.1547 | 1.51 | −0.42 | 10.90 |
| 4 | 3.92 | 188.27 | 21.47 | 0.76 | 0.6563 | 0.1935 | 1.77 | −0.01 | 10.25 |
| 5 | 4.71 | 142.47 | 19.27 | 0.87 | 0.6605 | 0.2456 | 2.05 | 0.43 | 9.77 |
| 6 | 5.63 | 108.56 | 16.95 | 0.97 | 0.6636 | 0.3242 | 2.37 | 0.82 | 8.62 |
| 7 | 6.80 | 80.43 | 14.53 | 1.04 | 0.6738 | 0.4501 | 2.76 | 1.48 | 6.71 |
| 8 | 8.40 | 58.69 | 12.21 | 1.12 | 0.7013 | 0.7067 | 3.27 | 2.34 | 3.75 |
| 9 | 11.01 | 39.92 | 9.57 | 1.15 | 0.7487 | 1.3002 | 4.07 | 4.01 | −0.85 |
| High MAX | 17.77 | 21.52 | 6.47 | 1.20 | 0.8890 | 4.0015 | 6.22 | 9.18 | −11.74 |

payoffs in the portfolio formation month, these stocks continue to exhibit this behavior in the future.

To get a clearer picture of the composition of the high MAX portfolios, Table 5 presents summary statistics for the stocks in the deciles. Specifically, the table reports the average across the months in the sample of the median values within each month of various characteristics for the stocks in each decile. We report values for the maximum daily return (in percent), the market capitalization (in millions of dollars), the price (in dollars), the market beta, the book-to-market (BM) ratio, a measure of illiquidity (scaled by $10^5$), the return in the portfolio formation month (REV), the return over the 11 months prior to portfolio formation (MOM), and the idiosyncratic volatility (IVOL).[7] Definitions of these variables are given in the Appendix.

The portfolios exhibit some striking patterns. As we move from the low MAX to the high MAX decile, the average across months of the median daily maximum return of stocks increases from 1.62% to 17.77%. With the exception of decile 10, these values are similar to those reported in Table 1 for the average maximum daily return. For decile 10, the average maximum return exceeds the median by approximately 6%. The distribution of maximum daily returns is clearly right-skewed, with some stocks exhibiting very high returns. These outliers are not a problem in the portfolio-level analysis, but we will revisit this issue in the firm-level, cross-sectional regressions.

As MAX increases across the deciles, market capitalization decreases. The absolute numbers are difficult to interpret since market capitalizations go up over time, but the relative values indicate that the high MAX portfolios are dominated by smaller stocks. This pattern is good news for the raw return differences shown in Table 1 since the concentration of small stocks in the high MAX deciles would suggest that these portfolios should earn a return premium, not the return discount observed in the data. This phenomenon may partially explain why the alpha difference exceeds the difference in raw returns.

The small stocks in the high MAX portfolios also tend to have low prices, declining to a median price of \$6.47 for decile 10. While this pattern is not surprising, it does suggest that there may be measurement issues associated with microstructure phenomena for some of the small, low-priced stocks in the higher MAX portfolios, or, more generally, that the results we show may be confined solely to micro-cap stocks with low stock prices. The fact that the results hold for value-weighted portfolios, as well as equal-weighted portfolios, does allay this concern somewhat, but it is still worthwhile to check the robustness of the results to different sample selection procedures.

First, we repeat the analysis in Table 1 excluding all stocks with prices below \$5/share. The four-factor alpha differences between the low MAX and high MAX value-weighted and equal-weighted portfolios are −0.81% and −1.14% per month, respectively, and both differences are highly statistically significant. Second, we exclude all Amex and Nasdaq stocks from the sample and form portfolios of stocks trading only on the NYSE. Again, the average risk-adjusted return differences are large and negative: −0.45% per month with a $t$-statistic of −2.48 for the value-weighted portfolios and −0.89% per month with a $t$-statistic of −5.15 for the equal-weighted portfolios. Finally, we sort all NYSE stocks by firm size each month to determine the NYSE decile breakpoints for market capitalization. Then, each month we exclude all NYSE/Amex/Nasdaq stocks with market capitalizations that would place them in the smallest NYSE size quintile, i.e., the two smallest size deciles, consistent with the definition of micro-cap stocks in Keim (1999) and Fama and French (2008). The average risk-adjusted return differences are −0.72% and −0.44% per month with $t$-statistics of −4.00 and −2.25 for the value-weighted and equal-weighted portfolios, respectively. These analyses provide convincing evidence that, while our main findings are certainly concentrated among smaller stocks,

[7] The qualitative results from the average statistics are very similar to those obtained from the median statistics. Since the median is a robust measure of the center of the distribution that is less sensitive to outliers than the mean, we choose to present the median statistics in Table 5.

the phenomenon is not confined to only the smallest, lowest-price segment of the market.

We can also look more directly at the distribution of market capitalizations within the high MAX decile. For example, during the last 2 years of our sample period, approximately 68% of these stocks fell below the size cutoff necessary for inclusion in the Russell 3000 index. In other words, almost one-third of the high MAX stocks were among the largest 3000 stocks. Over the full sample, approximately 50% of the high MAX stocks, on average, fell into the two smallest size deciles. This prevalence of small stocks with extreme positive returns, and their corresponding low future returns, is consistent with the theoretical motivation discussed earlier. It is individual investors, rather than institutions, that are most likely to be subject to the phenomena modeled in Barberis and Huang (2008) and Brunnermeier, Gollier, and Parker (2007), and individual investors also exhibit underdiversification. Thus, these effects should show up in the same small stocks that are held and traded by individual investors but by very few institutions.

Returning to the descriptive statistics in Table 5, betas are calculated monthly using a regression of daily excess stock returns on daily excess market returns; thus, these values are clearly noisy estimates of the true betas. Nevertheless, the monotonic increase in beta as MAX increases does suggest that stocks with high maximum daily returns are more exposed to market risk. To the extent that market risk explains the cross-section of expected returns, this relation between MAX and beta serves only to emphasize the low raw returns earned by the high MAX stocks as shown in Table 1. The difference in four-factor alphas should control for this effect, which partially explains why this difference is larger than the difference in the raw returns.

Median book-to-market ratios are similar across the portfolios, although if anything, high MAX portfolios do have a slight value tilt.

In contrast, the liquidity differences are substantial. Our measure of illiquidity is the absolute return over the month divided by the monthly trading volume, which captures the notion of price impact, i.e., the extent to which trading moves prices (see Amihud, 2002). We use monthly returns over monthly trading volume, rather than a monthly average of daily values of the same quantity, because a significant fraction of stocks have days with no trade. Eliminating these stocks from the sample reduces the sample size with little apparent change in the empirical results. Based on this monthly measure, illiquidity increases quite dramatically for the high MAX deciles, consistent with these portfolios containing smaller stocks. Again, this pattern only serves to strengthen the raw return differences shown in Table 1 since these stocks should earn a higher return to compensate for their illiquidity. Moreover, the four-factor alphas do not control for this effect except to the extent that the size and book-to-market factors also proxy for liquidity.

The final two columns of Table 5 report median returns in the portfolio formation month (REV) and the return over the previous 11 months (MOM). These two variables indicate the extent to which the portfolios are subject to short-term reversal and intermediate-term momentum effects, respectively. Jegadeesh and Titman (1993) and subsequent papers show that over intermediate horizons, stocks exhibit a continuation pattern, i.e., past winners continue to do well and past losers continue to perform badly. Over shorter horizons, stocks exhibit return reversals, due partly to microstructure effects such as bid-ask bounce (Jegadeesh, 1990; Lehmann, 1990).

The Fama-French-Carhart four-factor model does not control for short-term reversals; therefore, we control for the effects of REV in the context of bivariate sorts and cross-sectional regressions later in the paper. However, it is also possible that REV, a monthly return, does not adequately capture shorter-term effects. To verify that it is not daily or weekly microstructure effects that are driving our results, we subdivide the stocks in the high MAX portfolio according to when in the month the maximum daily return occurs. If the effect we find is more prominent for stocks whose maximum return occurs towards the end of the month, it would cast doubt on our interpretation of the evidence.

There is no evidence of this phenomenon. For example, for value-weighted portfolios, average raw return differences between the low MAX and high MAX portfolios are $-0.98\%$ per month for stocks with the maximum return in the first half of the month versus $-0.95\%$ per month for those with the maximum return in the second half of the month. The alpha differences follow the same pattern. Similarly, the raw return differences for stocks with the maximum return in the first week of the month are $-1.41\%$ per month, which is larger than the return difference of $-0.89\%$ per month for those stocks with maximum returns in the last week. Again, the alpha differences follow the same ordering. Moreover, the low returns associated with high MAX stocks persist beyond the first month after portfolio formation. Thus, short-term reversals at the daily or weekly frequency do not seem to explain the results.

Given that the portfolios are sorted on maximum daily returns, it is hardly surprising that median returns in the same month are also high, i.e., stocks with a high maximum daily return also have a high return that month. More interesting is the fact that the differences in median monthly returns for the portfolios of interest are smaller than the differences in the median MAX. For example, the difference in MAX between deciles 9 and 10 is 6.8% relative to a difference in monthly returns of 5.2%. In other words, the extreme daily returns on the lottery-like stocks are offset to some extent by lower returns on other days. This phenomenon explains why these same stocks can have lower average returns in the subsequent month (Table 1) even though they continue to exhibit a higher frequency of extreme positive returns (Tables 3 and 4).

This lower average return is also mirrored in the returns over the prior 11 months. The high MAX portfolios exhibit significantly lower and even negative returns over the period prior to the portfolio formation month. The strength of this relation is perhaps surprising, but it is consistent with the fact that stocks with extreme positive daily returns are small and have low prices.

The final column in Table 5 reports the idiosyncratic volatility of the MAX-sorted portfolios. It is clear that MAX and IVOL are strongly positively correlated in the cross-section. We address the relation between extreme returns and idiosyncratic volatility in detail in Section 3.

Given these differing characteristics, there is some concern that the four-factor model used in Table 1 to calculate alphas is not adequate to capture the true difference in risk and expected returns across the portfolios sorted on MAX. For example, the HML and SMB factors of Fama and French do not fully explain the returns of portfolios sorted by book-to-market ratios and size.[8] Moreover, the four-factor model does not control explicitly for the differences in expected returns due to differences in illiquidity or other known empirical phenomena such as short-term reversals. With the exception of short-term reversals and intermediate-term momentum, it seems unlikely that any of these factors can explain the return differences in Table 1 because high *MAX* stocks have characteristics that are usually associated with high expected returns, while these portfolios actually exhibit low returns. Nevertheless, in the following two subsections we provide different ways of dealing with the potential interaction of the maximum daily return with firm size, book-to-market, liquidity, and past returns. Specifically, we test whether the negative relation between *MAX* and the cross-section of expected returns still holds once we control for size, book-to-market, momentum, short-term reversal, and liquidity using bivariate portfolio sorts and Fama-MacBeth (1973) regressions.

### 2.3. *Bivariate portfolio-level analysis*

In this section we examine the relation between maximum daily returns and future stock returns after controlling for size, book-to-market, momentum, short-term reversals, and liquidity. For example, we control for size by first forming decile portfolios ranked based on market capitalization. Then, within each size decile, we sort stocks into decile portfolios ranked based on *MAX* so that decile 1 (decile 10) contains stocks with the lowest (highest) *MAX*. For brevity, we do not report returns for all 100 ($10 \times 10$) portfolios. Instead, the first column of Table 6, Panel A presents returns averaged across the ten size deciles to produce decile portfolios with dispersion in MAX, but which contain all sizes of firms. This procedure creates a set of MAX portfolios with similar levels of firm size, and thus, these MAX portfolios control for differences in size. After controlling for size, the value-weighted average return difference between the low MAX and high MAX portfolios is about $-1.22\%$ per month with a Newey–West $t$-statistic of $-4.49$. The 10–1 difference in the four-factor alphas is $-1.19\%$ per month with a $t$-statistic of $-5.98$. Thus, market capitalization does not explain the high (low) returns to low (high) MAX stocks.

If, instead of averaging across the size deciles, we look at the alpha differences for each decile in turn, the results are consistent with those reported in Section 2.2. Specifically, while the direction of the MAX effect is consistent across all the deciles, it is generally increasing in both magnitude and statistical significance as the market capitalization of the stocks decreases.

The fact that the results from the bivariate sort on size and MAX are, if anything, both economically and statistically more significant than those presented for the univariate sort in Table 1 is perhaps not too surprising. As shown in Table 5, the high MAX stocks, which have low subsequent returns, are generally small stocks. The standard size effect would suggest that these stocks should have high returns. Thus, controlling for size should enhance the effect on raw returns and even on four-factor alphas to the extent that the SMB factor is an imperfect proxy. However, there is a second effect of bivariate sorts that works in the opposite direction. Size and MAX are correlated; hence, variation in MAX within size-sorted portfolios is smaller than in the broader universe of stocks. That this smaller variation in MAX still generates substantial return variation is further evidence of the significance of this phenomenon.

The one concern with dependent bivariate sorts on correlated variables is that they do not sufficiently control for the control variable. In other words, there could be some residual variation in size across the MAX portfolios. We address this concern in two ways. First, we also try independent bivariate sorts on the two variables. These sorts produce very similar results. Second, in the next section we perform cross-sectional regressions in which all the variables appear as control variables.

We control for book-to-market (BM) in a similar way, with the results reported in the second column of Table 6, Panel A. Again the effect of MAX is preserved, with a value-weighted average raw return difference between the low MAX and high MAX deciles of $-0.93\%$ per month and a corresponding $t$-statistic of $-3.23$. The 10–1 difference in the four-factor alphas is also negative, $-1.06\%$ per month, and highly significant.

When controlling for momentum in column 3, the raw return and alpha differences are smaller in magnitude, but they are still economically large and statistically significant at all conventional levels. Again, the fact that momentum and MAX are correlated reduces the dispersion in maximum daily returns across the MAX portfolios, but intermediate-term continuation does not explain the phenomenon we show.

Column 4 controls for short-term reversals. Since firms with large positive daily returns also tend to have high monthly returns, it is conceivable that MAX could be proxying for the well-known reversal phenomenon at the monthly frequency, which we do not control for in the four-factor model in Table 1. However, this is not the case. After controlling for the magnitude of the monthly return in the portfolio formation month, the return and alpha differences are still 81 and 98 basis points, respectively, and both numbers exhibit strong statistical significance.

Finally, we control for liquidity by first forming decile portfolios ranked based on the illiquidity measure of

[8] Daniel and Titman (1997) attribute this failure to the fact that returns are driven by characteristics, not risk. We take no stand on this issue, but instead conduct a further battery of tests to demonstrate the robustness of our results.

**Table 6**
Returns on portfolios of stocks sorted by MAX after controlling for SIZE, BM, MOM, REV, and ILLIQ.

Double-sorted, value-weighted (Panel A) and equal-weighted (Panel B) decile portfolios are formed every month from July 1962 to December 2005 by sorting stocks based on the maximum daily returns after controlling for size, book-to-market, intermediate-term momentum, short-term reversals, and illiquidity. In each case, we first sort the stocks into deciles using the control variable, then within each decile, we sort stocks into decile portfolios based on the maximum daily returns over the previous month so that decile 1 (10) contains stocks with the lowest (highest) MAX. This table presents average returns across the ten control deciles to produce decile portfolios with dispersion in MAX but with similar levels of the control variable. "Return difference" is the difference in average monthly returns between the High MAX and Low MAX portfolios. "Alpha difference" is the difference in four-factor alphas on the High MAX and Low MAX portfolios. Newey-West (1987) adjusted $t$-statistics are reported in parentheses.

| *Panel A: Value-weighted portfolios* Decile | SIZE | BM | MOM | REV | ILLIQ |
|---|---|---|---|---|---|
| Low MAX | 1.47 | 1.22 | 1.32 | 1.06 | 1.29 |
| 2 | 1.60 | 1.19 | 1.14 | 1.18 | 1.31 |
| 3 | 1.69 | 1.27 | 1.17 | 1.19 | 1.30 |
| 4 | 1.65 | 1.19 | 1.07 | 1.18 | 1.23 |
| 5 | 1.57 | 1.17 | 1.03 | 1.15 | 1.12 |
| 6 | 1.49 | 1.23 | 1.03 | 1.15 | 1.06 |
| 7 | 1.29 | 1.13 | 0.96 | 1.04 | 0.99 |
| 8 | 1.20 | 0.99 | 0.93 | 1.07 | 0.88 |
| 9 | 0.93 | 0.89 | 0.88 | 0.86 | 0.60 |
| High MAX | 0.25 | 0.29 | 0.67 | 0.25 | 0.18 |
| Return difference | −1.22 | −0.93 | −0.65 | −0.81 | −1.11 |
| | (−4.49) | (3.23) | (−3.18) | (−2.70) | (−4.07) |
| Alpha difference | −1.19 | −1.06 | −0.70 | −0.98 | −1.12 |
| | (−5.98) | (−4.87) | (−5.30) | (−5.37) | (−5.74) |

| *Panel B: Equal-weighted portfolios* Decile | SIZE | BM | MOM | REV | ILLIQ |
|---|---|---|---|---|---|
| Low MAX | 1.52 | 1.37 | 1.47 | 1.36 | 1.40 |
| 2 | 1.63 | 1.50 | 1.45 | 1.56 | 1.59 |
| 3 | 1.73 | 1.53 | 1.38 | 1.60 | 1.60 |
| 4 | 1.70 | 1.54 | 1.32 | 1.58 | 1.58 |
| 5 | 1.62 | 1.48 | 1.29 | 1.59 | 1.52 |
| 6 | 1.54 | 1.52 | 1.20 | 1.53 | 1.52 |
| 7 | 1.38 | 1.45 | 1.15 | 1.44 | 1.40 |
| 8 | 1.27 | 1.33 | 1.08 | 1.33 | 1.32 |
| 9 | 1.04 | 1.19 | 1.03 | 1.15 | 1.05 |
| High MAX | 0.41 | 0.78 | 0.71 | 0.52 | 0.59 |
| Return difference | −1.11 | −0.59 | −0.76 | −0.83 | −0.81 |
| | (−4.05) | (−2.00) | (−3.70) | (−2.83) | (−2.68) |
| Alpha difference | −1.06 | −0.54 | −0.88 | −1.02 | −0.79 |
| | (−5.18) | (−1.96) | (−7.62) | (−5.09) | (−3.40) |

Amihud (2002), with the results reported in the final column of Table 6. Again, variation in MAX is apparently priced in the cross-section, with large return differences and corresponding $t$-statistics. Thus, liquidity does not explain the negative relation between maximum daily returns and future stock returns.

As mentioned earlier, we compute illiquidity as the ratio of the absolute monthly return to the monthly trading volume. We can also compute the original illiquidity measure of Amihud (2002), defined as the daily absolute return divided by daily dollar trading volume averaged within the month. These measures are strongly correlated, but in the latter case, we need to make a decision about how to handle stocks with zero trading volume on at least one day within the month. When we eliminate these stocks from the sample, the findings remain essentially unchanged. Raw return and alpha differences are −1.25% per month and −1.20% per month, respectively. Thus, for the remainder of the paper we focus on the larger sample and the monthly measure of illiquidity.

Next, we turn to an examination of the equal-weighted average raw and risk-adjusted returns on MAX portfolios after controlling for the same cross-sectional effects as in Table 6, Panel A. Again, to save space, instead of presenting the returns of all 100 ($10 \times 10$) portfolios for each control variable, we report the average returns of the MAX portfolios, averaged across the 10 control deciles to produce decile portfolios with dispersion in MAX but with similar levels of the control variable.

Table 6, Panel B shows that after controlling for size, book-to-market, momentum, short-term reversal, and liquidity, the equal-weighted average return differences between the low MAX and high MAX portfolios are −1.11%, −0.59%, −0.76%, −0.83%, and −0.81% per month, respectively. These average raw return differences are both economically and statistically significant. The corresponding values for the equal-weighted average risk-adjusted return differences are −1.06%, −0.54%, −0.88%, −1.02%, and −0.79%, which are also highly significant.

These results indicate that for both the value-weighted and the equal-weighted portfolios, the well-known cross-sectional effects such as size, book-to-market, momentum, short-term reversal, and liquidity cannot explain the low returns to high MAX stocks.

### 2.4. Firm-level cross-sectional regressions

So far we have tested the significance of the maximum daily return as a determinant of the cross-section of future returns at the portfolio level. This portfolio-level analysis has the advantage of being non-parametric in the sense that we do not impose a functional form on the relation between MAX and future returns. The portfolio-level analysis also has two potentially significant disadvantages. First, it throws away a large amount of information in the cross-section via aggregation. Second, it is a difficult setting in which to control for multiple effects or factors simultaneously. Consequently, we now examine the cross-sectional relation between MAX and expected returns at the firm level using Fama and MacBeth (1973) regressions.

We present the time-series averages of the slope coefficients from the regressions of stock returns on maximum daily return (MAX), market beta (BETA), log market capitalization (SIZE), log book-to-market ratio (BM), momentum (MOM), short-term reversal (REV), and illiquidity (ILLIQ). The average slopes provide standard Fama-MacBeth tests for determining which explanatory variables, on average, have non-zero premiums. Monthly cross-sectional regressions are run for the following econometric specification and nested versions thereof:

$$R_{i,t+1} = \lambda_{0,t} + \lambda_{1,t}\mathrm{MAX}_{i,t} + \lambda_{2,t}\mathrm{BETA}_{i,t} + \lambda_{3,t}\mathrm{SIZE}_{i,t} + \lambda_{4,t}\mathrm{BM}_{i,t} + \lambda_{5,t}\mathrm{MOM}_{i,t} + \lambda_{6,t}\mathrm{REV}_{i,t} + \lambda_{7,t}\mathrm{ILLIQ}_{i,t} + \varepsilon_{i,t+1}, \tag{1}$$

where $R_{i,t+1}$ is the realized return on stock $i$ in month $t$+1. The predictive cross-sectional regressions are run on the one-month lagged values of MAX, BETA, SIZE, BM, REV, and ILLIQ, and MOM is calculated over the 11-month period ending 2 months prior to the return of interest.

Table 7 reports the time-series averages of the slope coefficients $\lambda_{i,t}$ ($i$=1, 2, …, 7) over the 522 months from July 1962 to December 2005 for all NYSE/Amex/Nasdaq stocks. The Newey-West adjusted $t$-statistics are given in parentheses. The univariate regression results show a negative and statistically significant relation between the maximum daily return and the cross-section of future stock returns. The average slope, $\lambda_{1,t}$, from the monthly regressions of realized returns on MAX alone is −0.0434 with a $t$-statistic of −2.92. The economic magnitude of the associated effect is similar to that shown in Tables 1 and 6 for the univariate and bivariate sorts. The spread in median maximum daily returns between deciles 10 and 1 is approximately 16%. Multiplying this spread by the average slope yields an estimate of the monthly risk premium of −69 basis points.

In general, the coefficients on the individual control variables are also as expected—the size effect is negative and significant, the value effect is positive and significant, stocks exhibit intermediate-term momentum and short-term reversals, and illiquidity is priced. The average slope on BETA is negative and statistically insignificant, which contradicts the implications of the CAPM but is consistent with prior empirical evidence. In any case, these results should be interpreted with caution since BETA is estimated over a month using daily data, and thus, is subject to a significant amount of measurement error. The regression with all six control variables shows similar results, although the size effect is weaker and the coefficient on BETA is now positive, albeit statistically insignificant.

Of primary interest is the last line of Table 7, which shows the results for the full specification with MAX and the six control variables. In this specification, the average slope coefficient on MAX is −0.0637, substantially larger than in the univariate regression, with a commensurate increase in the $t$-statistic to −6.16. This coefficient corresponds to a 102 basis-point difference in expected monthly returns between median stocks in the high and low MAX deciles. The explanation for the increased magnitude of the estimated effect in the full specification is straightforward. Since stocks with high maximum daily returns tend to be small and illiquid, controlling for the increased expected return associated with these characteristics pushes the return premium associated with extreme positive return stocks even lower. These effects more than offset the reverse effect associated with intermediate-term momentum and short-term reversals, which partially explain the low future returns on high MAX stocks.

The strength of the results is somewhat surprising given that there are sure to be low-priced, thinly traded

**Table 7**
Firm-level cross-sectional return regressions.

Each month from July 1962 to December 2005 we run a firm-level cross-sectional regression of the return in that month on subsets of lagged predictor variables including MAX in the previous month and six control variables that are defined in the Appendix. In each row, the table reports the time-series averages of the cross-sectional regression slope coefficients and their associated Newey-West (1987) adjusted $t$-statistics (in parentheses).

| MAX | BETA | SIZE | BM | MOM | REV | ILLIQ |
|---|---|---|---|---|---|---|
| −0.0434 | | | | | | |
| (−2.92) | | | | | | |
| | −0.0118 | | | | | |
| | (−0.43) | | | | | |
| | | −0.1988 | | | | |
| | | (−4.08) | | | | |
| | | | 0.4651 | | | |
| | | | (6.73) | | | |
| | | | | 0.7317 | | |
| | | | | (4.67) | | |
| | | | | | −0.0675 | |
| | | | | | (−11.24) | |
| | | | | | | 0.0371 |
| | | | | | | (3.87) |
| | 0.0140 | −0.0865 | 0.3390 | 0.7436 | −0.0751 | 0.0223 |
| | (0.56) | (−1.73) | (4.82) | (5.29) | (−14.15) | (3.64) |
| −0.0637 | 0.0485 | −0.1358 | 0.3201 | 0.6866 | −0.0712 | 0.0224 |
| (−6.16) | (2.18) | (−3.10) | (4.69) | (4.97) | (−13.53) | (3.78) |

stocks within our sample whose daily returns will exhibit noise due to microstructure and other effects. To confirm this intuition, we re-run the cross-sectional regressions after winsorizing MAX at the 99th and 95th percentiles to eliminate outliers. In the full specification, the average coefficient on MAX increases to −0.0788 and −0.0902, suggesting that the true economic effect is even larger than that shown in Table 7. A different but related robustness check is to run the same analysis using only NYSE stocks, which tend to be larger and more actively traded and are thus likely to have less noisy daily returns. For this sample, the baseline coefficient of −0.064 in Table 7 increases to −0.077.

Given the characteristics of the high MAX stocks, as discussed previously, it is also worthwhile verifying that different methods of controlling for illiquidity do not affect the main results. Using the daily Amihud (2002) measure averaged over the month, the coefficient on MAX is somewhat larger in magnitude. In addition, controlling for the liquidity risk measure of Pastor and Stambaugh (2003) has little effect on the results.

The regression in Eq. (1) imposes a linear relation between returns and MAX for simplicity rather than for theoretical reasons. However, adding a quadratic term to the regression or using a piecewise linear specification appears to add little, if anything, to the explanatory power. Similarly, interacting MAX with contemporaneous volume, with the idea that trading volume may be related to the informativeness of the price movements, also proved fruitless.

The clear conclusion is that cross-sectional regressions provide strong corroborating evidence for an economically and statistically significant negative relation between extreme positive returns and future returns, consistent with models that suggest that idiosyncratic lottery-like payoffs are priced in equilibrium.

## 3. Idiosyncratic volatility and extreme returns

While arguably MAX is a theoretically motivated variable, there is still a concern that it may be proxying for a different effect. In particular, stocks with high volatility are likely to exhibit extreme returns of both signs. Moreover, stocks with high maximum daily returns in a given month will also have high realized volatility in the same month, measured using squared daily returns, almost by construction. Ang, Hodrick, Xing, and Zhang (2006, 2009) show that idiosyncratic volatility has a significant negative price in the cross-section, i.e., stocks with high idiosyncratic volatility have low subsequent returns[9]; thus, it is plausible that MAX is proxying for this effect. We examine this issue in detail in this section.

**Table 8**
Time-series average of cross-sectional correlations.

The table reports the average across months of the cross-sectional correlation of the maximum daily return in a month (MAX), the average of the highest five daily returns in a month (MAX(5)), the minimum daily return in a month (MIN), total volatility (TVOL), and idiosyncratic volatility (IVOL) for the period July 1962 to December 2005.

| | MAX | MAX(5) | MIN | TVOL | IVOL |
|---|---|---|---|---|---|
| MAX | 1 | 0.8981 | 0.5491 | 0.7591 | 0.7533 |
| MAX(5) | | 1 | 0.6153 | 0.8312 | 0.8204 |
| MIN | | | 1 | 0.7603 | 0.7554 |
| TVOL | | | | 1 | 0.9842 |
| IVOL | | | | | 1 |

As preliminary evidence, Table 8 provides the average monthly cross-sectional correlations between five variables of interest—MAX (the maximum daily return within the month), MAX(5) (the average of the highest five daily returns within the month), MIN (the negative of the minimum daily return within the month), TVOL (monthly realized total volatility measured using daily returns within the month), and IVOL (monthly realized idiosyncratic volatility measured using the residuals from a daily market model within the month). TVOL, IVOL, and MIN are defined in the Appendix. We reverse the sign on the minimum daily returns so that high values of MIN correspond to more extreme returns. Note that idiosyncratic volatility and total volatility are essentially identical when measured within a month due to the low explanatory power of the market model regression. In our sample, the average cross-sectional correlation between these variables exceeds 0.98. We choose to work with IVOL since it corresponds to the variable used by Ang, Hodrick, Xing, and Zhang (2006).[10]

Not surprisingly, MAX and MAX(5) are highly correlated. Of greater interest, the average, cross-sectional correlations between IVOL and both MAX and MIN are approximately 0.75, which is very high given that all three variables are calculated at the individual stock level. MAX(5) is even more highly correlated with IVOL than MAX. Moreover, these correlations are not driven simply by the fact that a squared extreme daily return leads to a high measured realized volatility. Even when the maximum and minimum daily returns are eliminated prior to the calculation of volatility, volatility remains highly correlated with MAX, MAX(5), and MIN. MAX and MAX(5) are also quite closely related to MIN, with correlations of 0.55 and 0.62, respectively. Clearly stocks with high volatility exhibit extreme returns and vice versa.

A second important piece of preliminary evidence is to verify the relation between idiosyncratic volatility and

[9] Fu (2009) emphasizes the time-series variation in idiosyncratic volatility and finds a significantly positive relation between conditional idiosyncratic variance and the cross-section of expected returns. Spiegel and Wang (2005) estimate idiosyncratic volatility from monthly rather than daily returns and find that stock returns increase with the level of idiosyncratic risk and decrease with the stock's liquidity but that idiosyncratic risk often subsumes the explanatory power of liquidity. Fu (2009) and Huang, Liu, Rhee, and Zhang (2010) argue that the results are driven by monthly stock-return reversals, although Nyberg (2008) disputes this claim. Fang and Peress (2009) show that the idiosyncratic volatility effect is reversed for stocks with no media coverage.

[10] Measuring idiosyncratic volatility relative to a three-factor or four-factor model rather than the market model has little effect on the results.

future returns in our sample. We conduct a univariate portfolio sort on IVOL, similar to that given in Table 1 for MAX, although, for brevity, we do not report the results in detail. These results look very similar to those in Table 1. For value-weighted returns, deciles 1 through 7 (lower idiosyncratic volatility) all exhibit average monthly returns of around 1%. These returns fall dramatically for the higher volatility stocks, all the way to 0.02% per month for decile 10. Both the return differences between the low and high IVOL deciles of −0.93% per month and the corresponding four-factor alpha differences of −1.33% are economically and statistically significant. These results coincide closely with the results in Ang, Hodrick, Xing, and Zhang (2006), although they form quintiles rather than deciles and use a slightly shorter sample period. Of some interest, there is no evidence of an idiosyncratic volatility effect in equal-weighted portfolios—a result that is found in Bali and Cakici (2008). Given the strong positive correlation between MAX and IVOL shown in Table 8 above, it is not surprising that average maximum daily returns increase across the IVOL-sorted portfolios. In fact, the range is not that much smaller than in the MAX-sorted portfolios (Table 1), with the stocks in the low and high IVOL portfolios having an average MAX of 1.95% and 17.31%, respectively.

To examine the relation between extreme returns and volatility more closely, we first conduct four bivariate sorts. In Table 9, Panel A we sort on both the maximum daily return (MAX) and the average of the five highest daily returns (MAX(5)), controlling for idiosyncratic volatility. We first form decile portfolios ranked based

**Table 9**

Returns on portfolios of stocks sorted by MAX and IVOL after controlling for IVOL and MAX.

Double-sorted, value-weighted (VW) and equal-weighted (EW) decile portfolios are formed every month from July 1962 to December 2005. In Panel A we sort stocks based on the maximum daily return (MAX) or average of the five highest daily returns (MAX(5)) after controlling for idiosyncratic volatility (IVOL). In Panel B we sort stocks based on idiosyncratic volatility (IVOL) after controlling for the maximum daily return (MAX) or average of the five highest daily returns (MAX(5)). In both cases, we first sort the stocks into deciles using the control variable, then within each decile, we sort stocks into decile portfolios based on the variable of interest. The columns report average returns across the ten control deciles to produce decile portfolios with dispersion in the variable of interest but with similar levels of the control variable. "Return difference" is the difference in average monthly returns between deciles 10 and 1. "Alpha difference" is the difference in four-factor alphas between deciles 10 and 1. Newey-West (1987) adjusted *t*-statistics are reported in parentheses.

*Panel A: Sorted by MAX and MAX(5) controlling for IVOL*

| | $N=1$ | | $N=5$ | |
|---|---|---|---|---|
| Decile | VW | EW | VW | EW |
| Low MAX($N$) | 1.12 | 2.01 | 1.39 | 2.25 |
| 2 | 1.09 | 1.65 | 1.18 | 1.81 |
| 3 | 0.94 | 1.54 | 1.20 | 1.67 |
| 4 | 0.93 | 1.41 | 1.11 | 1.51 |
| 5 | 0.80 | 1.34 | 0.99 | 1.38 |
| 6 | 0.77 | 1.22 | 0.84 | 1.21 |
| 7 | 0.79 | 1.19 | 0.74 | 1.11 |
| 8 | 0.82 | 1.23 | 0.79 | 1.06 |
| 9 | 0.76 | 1.04 | 0.67 | 0.93 |
| High MAX($N$) | 0.77 | 1.10 | 0.53 | 0.75 |
| Return difference | −0.35 | −0.91 | −0.86 | −1.50 |
| | (−2.42) | (−7.86) | (−4.36) | (−9.21) |
| Alpha difference | −0.34 | −0.92 | −0.84 | −1.58 |
| | (−2.48) | (−7.96) | (−4.98) | (−10.05) |

*Panel B: Sorted by IVOL controlling for MAX and MAX(5)*

| | MAX | | MAX(5) | |
|---|---|---|---|---|
| Decile | VW | EW | VW | EW |
| Low IVOL | 1.03 | 1.18 | 0.89 | 0.84 |
| 2 | 0.93 | 1.15 | 0.86 | 1.02 |
| 3 | 0.90 | 1.10 | 0.78 | 1.03 |
| 4 | 0.92 | 1.17 | 0.93 | 1.17 |
| 5 | 0.95 | 1.27 | 0.97 | 1.20 |
| 6 | 0.88 | 1.21 | 0.98 | 1.28 |
| 7 | 0.94 | 1.37 | 0.99 | 1.40 |
| 8 | 0.83 | 1.48 | 1.09 | 1.56 |
| 9 | 0.73 | 1.52 | 0.96 | 1.69 |
| High IVOL | 0.66 | 2.16 | 0.95 | 2.51 |
| Return difference | −0.38 | 0.98 | 0.06 | 1.67 |
| | (−1.98) | (4.88) | (0.29) | (8.04) |
| Alpha difference | −0.44 | 0.95 | 0.05 | 1.74 |
| | (−3.12) | (4.76) | (0.34) | (7.67) |

on idiosyncratic volatility, and within each IVOL decile we sort stocks into decile portfolios based on MAX or MAX(5) so that decile 1 (decile 10) contains stocks with the lowest (highest) MAX(N). Panel A shows the average of the value-weighted and equal-weighted returns across the IVOL deciles and the associated Newey-West *t*-statistics. The key statistics are the return and four-factor alpha differences (and Newey-West *t*-statistics) between the low MAX(N) and high MAX(N) portfolios, i.e., the differences between returns on portfolios that vary in MAX(N) but have approximately the same levels of idiosyncratic volatility.

The value-weighted average raw return difference between the low MAX and high MAX deciles is $-0.35\%$ per month with a *t*-statistic of $-2.42$. The 10–1 difference in the four-factor alphas is also negative, $-0.34\%$ per month, and highly significant. These magnitudes are much smaller than we have seen previously, but this result is hardly surprising. Idiosyncratic volatility and MAX are highly correlated; thus, after controlling for idiosyncratic volatility, the spread in maximum returns is significantly reduced. Nevertheless, idiosyncratic volatility does not completely explain the high (low) returns to low (high) MAX stocks. The equal-weighted average raw and risk-adjusted return differences between the low MAX and high MAX portfolios are much more negative, greater than 90 basis points per month in absolute magnitude, and highly significant with the *t*-statistics of $-7.86$ to $-7.96$, respectively. However, recall that the idiosyncratic volatility effect does not exist in equal-weighted portfolios.

When we sort on the average of the five highest daily returns within the month, the return and alpha differences for both value-weighted and equal-weighted portfolios exhibit substantially greater economic and statistical significance, consistent with the univariate results reported in Table 2. In both cases, if we examine the alpha differences individually for each IVOL decile, the pattern is intuitive. Given the high correlation between MAX and IVOL, it is only in the higher IVOL deciles where there are larger numbers of stocks with extreme positive returns. Therefore, the MAX effect tends to increase in magnitude and statistical significance as IVOL increases.

What happens if we perform the reverse sort, i.e., if we examine the explanatory power of idiosyncratic volatility after controlling for MAX(N)? In Table 9, Panel B, we first form decile portfolios ranked based either on the maximum daily returns over the past one month (MAX) or the average of the five highest daily returns (MAX(5)). Then, within each MAX(N) decile, we sort stocks into decile portfolios ranked based on IVOL so that decile 1 (decile 10) contains stocks with the lowest (highest) IVOL. When controlling for MAX, the average value-weighted raw return difference between the low IVOL and high IVOL portfolios is $-0.38\%$ per month with a *t*-statistic of $-1.98$. The 10–1 difference in the four-factor alphas is also negative, $-0.44\%$ per month, and statistically significant. These magnitudes are much smaller than those obtained from the univariate volatility portfolios; nevertheless, for the value-weighted portfolios, maximum daily return does not completely explain the idiosyncratic volatility puzzle in a simple bivariate sort.

There are two possible explanations for this result in combination with the results of Table 9, Panel A, and the significance of IVOL in the context of a univariate sort. First, MAX and IVOL could be picking up separate effects, both of which exist in the data. The absence of an idiosyncratic volatility effect in equal-weighted portfolios could be due to measurement issues for smaller stocks. Alternatively, it could be that bivariate sorts are not powerful enough to disentangle the true effect. While the idea of the bivariate sort is to produce portfolios with variation in the variable of interest but similar levels of the control variable, this goal is extremely difficult to achieve for highly correlated variables. While the stocks in the portfolios whose returns are reported in the first column of Table 9, Panel B do vary in their levels of idiosyncratic volatility, they also vary in their maximum daily returns. For example, the averages of the median idiosyncratic volatilities are 1.69% and 4.57% for the low and high IVOL portfolios, respectively, but the averages of the median MAX for these portfolios are 6.03% and 8.90%. Thus, it is difficult to know which effect is actually producing the negative return and alpha differences between these portfolios. One might think that an independent bivariate sort would solve this problem. Unfortunately, such a sort is infeasible because there are so few stocks with extreme positive returns and low volatility, or high volatility and no extreme returns. As a result, the portfolios of interest are exactly those for which we cannot observe reliable returns.

However, columns 2–4 of Panel B do shed further light on the issue of disentangling the effects of the two variables. In column 2, we report the results for equal-weighted portfolios, controlling for MAX. The average return difference between the high IVOL and low IVOL portfolios is about 0.98% per month with a Newey-West *t*-statistic of 4.88. The 10–1 difference in the four-factor alphas is 0.95% per month with a *t*-statistic of 4.76. Thus, after controlling for MAX, we find a significant and positive relation between IVOL and the cross-section of expected returns. This is the reverse of the counter-intuitive negative relation shown by Ang, Hodrick, Xing, and Zhang (2006, 2009). Once we control for extreme positive returns, there appears to be a reward for holding idiosyncratic risk. This result is consistent with a world in which risk-averse and poorly diversified agents set prices, yet these agents have a preference for lottery-like assets, i.e., assets with extreme positive returns in some states.

First, note that measurement error in idiosyncratic volatility cannot explain this positive and significant relation between idiosyncratic volatility and returns. Measurement error in the sorting variable will push return differences toward zero, but it cannot explain a sign reversal that is statistically significant, especially at the levels we report. Second, the inability to adequately control for variation in the control variable MAX is also not a viable explanation for these results. Residual variation in MAX is generating, if anything, the opposite effect. Finally, a positive relation between idiosyncratic volatility and returns and a negative relation between MAX and returns provides an explanation for the absence of a univariate idiosyncratic volatility effect in equal-weighted portfolios. This particular weighting scheme causes the IVOL and MAX effects to cancel, generating small and insignificant return differences.

To confirm these conclusions, the last two columns of Table 9, Panel B present results for portfolios that control for our somewhat more powerful measure of extreme returns, the average of the five highest daily returns during the month (MAX(5)). Using this control variable, the differences between the raw and risk-adjusted returns on high IVOL and low IVOL portfolios are positive, albeit insignificant, and the differences for equal-weighted portfolios are positive and extremely economically and statistically significant. The evidence supports the theoretically coherent hypothesis that lottery-like stocks command a price premium and those with high idiosyncratic risk trade at a discount.

We further examine the cross-sectional relation between IVOL and expected returns at the firm level using Fama-MacBeth regressions, with the results reported in the top half of Table 10. In the univariate regression, the average slope coefficient on IVOL is negative, −0.05, but it is not statistically significant ($t$-stat = −0.97). This lack of significance mirrors the result that there is little or no relation between volatility and future returns in equal-weighted portfolios. The cross-sectional regressions put equal weight on each firm observation.

When we add MAX to the regression, the negative relation between idiosyncratic volatility and expected returns is reversed. Specifically, the estimated average slope coefficient on IVOL is 0.39 with a Newey-West $t$-statistic of 4.69. This positive relation between IVOL and expected returns remains significant even after augmenting the regression with the six control variables.

Based on the bivariate equal-weighted portfolios and the firm-level cross-sectional regressions with MAX and IVOL, our conclusion is that there is no idiosyncratic volatility puzzle as recently reported in Ang, Hodrick, Xing, and Zhang (2006, 2009). In fact, stocks with high idiosyncratic volatility have higher future returns as would be expected in a world where poorly diversified and risk-averse investors help determine prices. We conclude that the reason for the presence of a negative relation between IVOL and expected returns shown by Ang et al. is that IVOL is a proxy for MAX. Interestingly, Han and Kumar (2008) provide evidence that the idiosyncratic volatility puzzle is concentrated in stocks dominated by retail investors. This evidence complements our results, since it is retail investors who are more likely to suffer from underdiversification and exhibit a preference for lottery-like assets.

A slightly different way to examine the relation between extreme returns and volatility is to look at minimum returns. If it is a volatility effect that is driving returns, then MIN (the minimum daily return over the month), which is also highly correlated with volatility, should generate a similar effect to MAX. On the other hand, much of the theoretical literature would predict that the effect of MIN should be the opposite of that of MAX. For example, if investors have a skewness preference, then stocks with negatively skewed returns should require higher returns. Similarly, under the cumulative prospect theory of Barberis and Huang (2008), small probabilities or large losses are over-weighted, and thus, these stocks have lower prices and higher expected returns.

To examine this issue, we form portfolios of stocks sorted on MIN after controlling for MAX. For brevity the results are not reported, but the return and alpha differences are positive and statistically significant,

**Table 10**
Firm-level cross-sectional return regressions with MAX, MIN, and IVOL.
Each month from July 1962 to December 2005 we run a firm-level cross-sectional regression of the return in that month on subsets of lagged predictor variables including MAX, MIN, and IVOL in the previous month and six control variables that are defined in the Appendix. In each row, the table reports the time-series averages of the cross-sectional regression slope coefficients and their associated Newey-West (1987) adjusted $t$-statistics (in parentheses).

| MAX | IVOL | MIN | BETA | SIZE | BM | MOM | REV | ILLIQ |
|---|---|---|---|---|---|---|---|---|
| −0.0434 | | | | | | | | |
| (−2.92) | | | | | | | | |
| | −0.0530 | | | | | | | |
| | (−0.97) | | | | | | | |
| −0.1549 | 0.3857 | | | | | | | |
| (−10.19) | (4.69) | | | | | | | |
| −0.0961 | 0.1210 | | 0.0496 | −0.1047 | 0.3244 | 0.7199 | −0.0716 | 0.0232 |
| (−7.90) | (1.95) | | (2.27) | (−2.71) | (4.85) | (5.35) | (−14.27) | (3.76) |
| | | 0.0593 | | | | | | |
| | | (2.41) | | | | | | |
| −0.0900 | | 0.1280 | | | | | | |
| (−7.84) | | (6.21) | | | | | | |
| −0.0761 | | 0.0393 | 0.0390 | −0.1135 | 0.3312 | 0.7098 | −0.0694 | 0.0223 |
| (−8.62) | | (2.66) | (1.87) | (−2.77) | (4.95) | (5.18) | (−14.35) | (3.74) |
| −0.1103 | 0.0840 | 0.1029 | | | | | | |
| (−6.90) | (0.94) | (5.43) | | | | | | |
| −0.0886 | 0.0589 | 0.0219 | 0.0382 | −0.1082 | 0.3287 | 0.7100 | −0.0706 | 0.0230 |
| (−6.67) | (0.80) | (1.45) | (1.80) | (−2.80) | (4.92) | (5.31) | (−14.75) | (3.73) |

although both the magnitudes and levels of significance are lower than those for MAX. This evidence suggests that stocks with extreme low returns have higher expected returns in the subsequent month. The opposite effects of MAX and MIN are consistent with cumulative prospect theory and skewness preference, but they are not consistent with the hypothesis that extreme returns are simply proxying for idiosyncratic volatility.

In addition to the portfolio-level analyses, we run firm-level Fama-MacBeth cross-sectional regressions with MAX, MIN, and IVOL. The bottom half of Table 10 presents the average slope coefficients and the Newey-West adjusted *t*-statistics. For all econometric specifications, the average slope on MAX remains negative and significant, confirming our earlier findings from the bivariate sorts. After controlling for MIN and IVOL, as well as market beta, size, book-to-market, momentum, short-term reversals, and liquidity, the average slope on MAX is $-0.089$ with a *t*-statistic of $-6.67$.

For specifications with MAX and MIN, but not IVOL, the average slope on MIN is positive and both economically and statistically significant. Note that the original minimum returns are multiplied by $-1$ in constructing the variable MIN. Therefore, the positive slope coefficient means that the more a stock fell in value, the higher the future expected return. The addition of the six control variables clearly weakens the estimated effect. This result is not surprising since stocks with extreme negative returns have characteristics similar to those of firms with extreme positive returns, i.e., they tend to be small and illiquid. Thus, size and illiquidity both serve to explain some of the positive returns earned by these stocks. Moreover, the MIN effect is not robust to the same subsampling exercises we report in Section 2.2 for MAX. When we exclude stocks whose market capitalizations would place them in the smallest NYSE size quintile, or when we examine NYSE stocks only, the MIN effect is no longer statistically significant, and, in fact, the sign of the effect is often reversed. Thus, the MIN effect, in contrast to the MAX effect, appears to be limited to micro-cap stocks. This result is perhaps not surprising because it may be costly to engage in the short-selling necessary to exploit the MAX effect, while exploiting the MIN effect involves taking a long position in the relevant stocks.

For the full specification with MAX, MIN, and IVOL, the coefficients on MIN and IVOL are no longer statistically significant. However, this result is most likely due to the multicollinearity in the regression, i.e., the correlations between MIN and IVOL (see Table 8) and between MIN, IVOL, and the control variables. The true economic effect of extreme negative returns is still an open issue, but these regressions provide further evidence that there is no idiosyncratic volatility puzzle.

## 4. Skewness and MAX

Our final empirical exercise is to examine the link, if any, between extreme positive returns and skewness in terms of their ability to explain the cross-section of expected returns. The investigation of the role of higher moments in asset pricing has a long history. Arditti (1967), Kraus and Litzenberger (1976), and Kane (1982) extend the standard mean-variance portfolio theory to incorporate the effect of skewness on valuation. They present a three-moment asset pricing model in which investors hold concave preferences and like positive skewness. In this framework, assets that decrease a portfolio's skewness (i.e., that make the portfolio returns more left-skewed) are less desirable and should command higher expected returns. Similarly, assets that increase a portfolio's skewness should generate lower expected returns.[11]

From our perspective, the key implication of these models is that it is systematic skewness, not idiosyncratic skewness, that explains the cross-sectional variation in stock returns. Investors hold the market portfolio in which idiosyncratic skewness is diversified away, and thus, the appropriate measure of risk is co-skewness—the extent to which the return on an individual asset covaries with the variance of market returns. Harvey and Siddique (1999, 2000) and Smith (2007) measure conditional co-skewness and find that stocks with lower co-skewness outperform stocks with higher co-skewness, consistent with the theory, and that this premium varies significantly over time.

In contrast, the extreme daily returns measured by MAX are almost exclusively idiosyncratic in nature, at least for the high MAX stocks, which produce the anomalous, low subsequent returns. Of course, this does not mean that MAX is not proxying for the systematic skewness, or co-skewness, of stocks. Thus, the first question is whether MAX, despite its idiosyncratic nature, is robust to controls for co-skewness.

The second question is whether MAX is priced because it proxies for idiosyncratic skewness. In other words, is MAX simply a good proxy for the third moment of returns? There is some empirical evidence for a skewness effect in returns. For example, Zhang (2005) computes a measure of cross-sectional skewness, e.g., the skewness of firm returns within an industry, that predicts future returns at the portfolio level. Boyer, Mitton, and Vorkink (2010) employ a measure of expected skewness, i.e., a projection of 5-year-ahead skewness on a set of predetermined variables, including stock characteristics, to predict portfolio returns over the subsequent month. Finally, Conrad, Dittmar, and Ghysels (2008) show that measures of risk-neutral skewness from option prices predict subsequent returns. In all three cases, the direction of the results is consistent with our evidence, i.e., more positively skewed stocks have lower returns, but these effects are generally weaker than the economically and statistically strong evidence we provide in Section 2.

Of equal importance, there is no theoretical reason to prefer return skewness to extreme returns as a potential variable to explain the cross-section of expected returns. In the model of Barberis and Huang (2008), based on the cumulative prospect theory of Tversky and Kahneman

[11] Arditti (1971), Friend and Westerfield (1980), Sears and Wei (1985), Barone-Adesi (1985), and Lim (1989) provide empirical analyses of the role of skewness.

(1992), it is the low probability, extreme return states that drive the results, not skewness directly. Similarly, in the optimal beliefs model of Brunnermeier, Gollier, and Parker (2007), it is again low probability states that drive the relevant pricing effects. Only in the model of Mitton and Vorkink (2007), who assume a preference for positive skewness, is skewness the natural measure.

To determine whether the information content of maximum daily returns and skewness is similar, we test the significance of the cross-sectional relation between MAX and future stock returns after controlling for total skewness (TSKEW), idiosyncratic skewness (ISKEW), and systematic skewness (SSKEW). In contrast with our other control variables, we calculate these skewness measures primarily over one year using daily returns.[12] A one-year horizon provides a reasonable tradeoff between having a sufficient number of observations to estimate skewness and accommodating time-variation in skewness. Total skewness is the natural measure of the third central moment of returns; systematic skewness, or co-skewness, is the coefficient of a regression of returns on squared market returns, including the market return as a second regressor (as in Harvey and Siddique, 2000); and idiosyncratic skewness is the skewness of the residuals from this regression. These variables are defined in more detail in the Appendix. Total skewness and idiosyncratic skewness are similar for most stocks due to the low explanatory power of the regression using daily data.

We first perform bivariate sorts on MAX while controlling for skewness. We control for total skewness by forming decile portfolios ranked based on TSKEW. Then, within each TSKEW decile, we sort stocks into decile portfolios ranked based on MAX so that decile 1 (decile 10) contains stocks with the lowest (highest) MAX. The first column of Table 11 shows returns averaged across the 10 TSKEW deciles to produce decile portfolios with dispersion in MAX, but which contain firms with all levels of total skewness. After controlling for total skewness, the value-weighted average return difference between the low MAX and high MAX portfolios is about −0.94% per month with a Newey-West *t*-statistic of −3.06. The 10–1 difference in the four-factor alphas is −1.00% per month with a *t*-statistic of −4.34. Thus, total skewness does not explain the high (low) returns to low (high) MAX stocks.

The last two columns of Table 11 present similar results from the bivariate sorts of portfolios formed based on MAX after controlling for systematic and idiosyncratic skewness, respectively. After controlling for systematic skewness, or co-skewness, the value-weighted average raw and risk-adjusted return differences between the low MAX and high MAX portfolios are in the range of 110–123 basis points per month and highly significant. After controlling for idiosyncratic skewness, the value-weighted average raw and risk-adjusted return differences between the low MAX and high MAX portfolios are −0.93% to −1.01% per month with the *t*-statistics of −2.96 and −4.34, respectively. These results indicate that systematic and idiosyncratic skewness

**Table 11**
Returns on portfolios of stocks sorted by MAX after controlling for skewness.

Double-sorted, value-weighted decile portfolios are formed every month from July 1962 to December 2005 by sorting stocks based on the maximum daily returns after controlling for total (TSKEW), systematic (SSKEW), and idiosyncratic skewness (ISKEW). In each case, we first sort the stocks into deciles using the control variable, then within each decile, we sort stocks into decile portfolios based on the maximum daily returns over the previous month so that decile 1 (10) contains stocks with the lowest (highest) MAX. The table reports average returns across the ten control deciles to produce decile portfolios with dispersion in MAX but with similar levels of the control variable. "Return difference" is the difference in average monthly returns between high MAX and low MAX portfolios. "Alpha difference" is the difference in four-factor alphas between high MAX and low MAX portfolios. Newey-West (1987) adjusted *t*-statistics are reported in parentheses.

| Decile | TSKEW | SSKEW | ISKEW |
|---|---|---|---|
| Low MAX | 1.06 | 1.12 | 1.04 |
| 2 | 1.11 | 1.06 | 1.14 |
| 3 | 1.21 | 1.06 | 1.18 |
| 4 | 1.07 | 1.10 | 1.08 |
| 5 | 1.13 | 1.11 | 1.17 |
| 6 | 1.14 | 1.10 | 1.10 |
| 7 | 0.97 | 0.98 | 0.99 |
| 8 | 0.87 | 0.89 | 0.91 |
| 9 | 0.76 | 0.80 | 0.74 |
| High MAX | 0.12 | 0.03 | 0.11 |
| Return difference | −0.94 | −1.10 | −0.93 |
| | (−3.06) | (−3.75) | (−2.96) |
| Alpha difference | −1.00 | −1.23 | −1.01 |
| | (−4.34) | (−5.50) | (−4.34) |

cannot explain the significantly negative relation between MAX and expected stock returns.

One concern with this analysis is that lagged skewness may not be a good predictor of future skewness, as argued by Boyer, Mitton, and Vorkink (2010). In a rational market, it is expected future skewness that matters. This issue is addressed in Table 12, which presents results from cross-sectional, firm-level regressions of total skewness on lagged values of total skewness and our six control variables.[13] Skewness is significantly persistent, both in a univariate and multivariate context, although the explanatory power of the regressions is not very high. One possibility is to use the fitted values from the month-by-month cross-sectional regressions as a measure of expected skewness (as in Boyer, Mitton, and Vorkink, 2010), and thus, we include this variable in the cross-sectional return regressions that follow.

Table 13 presents the cross-sectional Fama-MacBeth regression results including TSKEW, SSKEW, ISKEW, and expected total skewness (E(TSKEW)) as control variables. The table reports the time-series averages of the slope coefficients over the sample period July 1962–December 2005, with Newey-West adjusted *t*-statistics given in parentheses. The inclusion of any of the skewness measures has only a limited effect on MAX. The average coefficients on MAX in the different specifications are all

[12] We test the robustness of our conclusions to variation in the measurement horizon (one, 3, 6, and 12 months) and find similar results.

[13] Using idiosyncratic skewness generates similar results.

**Table 12**
Cross-sectional predictability of skewness.

Each month from July 1962 to December 2005 we run a firm-level cross-sectional regression of the total skewness measured using daily returns over the subsequent year (TSKEW) on subsets of lagged predictor variables including TSKEW in the previous year and six control variables that are defined in the Appendix. The table reports the time-series averages of the cross-sectional regression coefficients, their associated Newey-West (1987) adjusted *t*-statistics (in parentheses), and the regression *R*-squareds.

| TSKEW | BETA | SIZE | BM | MOM | REV | ILLIQ | $R^2$ |
|---|---|---|---|---|---|---|---|
| 0.1507 | | | | | | | 2.46% |
| (28.38) | | | | | | | |
| 0.0813 | 0.0026 | −0.1218 | −0.0472 | 0.1489 | 0.0007 | 0.0271 | 9.69% |
| (21.70) | (1.82) | (−23.10) | (−6.30) | (12.70) | (3.79) | (1.94) | |

**Table 13**
Firm-level cross-sectional return regressions with MAX and skewness.

Each month from July 1962 to December 2005 we run a firm-level cross-sectional regression of the return in that month on subsets of lagged predictor variables including MAX in the previous month, skewness measured over the preceding year (TSKEW, SSKEW, ISKEW), fitted expected total skewness (*E*(TSKEW)) based on the regression in Table 12, and six control variables that are defined in the Appendix. In each row, the table reports the time-series averages of the cross-sectional regression slope coefficients and their associated Newey-West (1987) adjusted *t*-statistics (in parentheses).

| MAX | BETA | SIZE | BM | MOM | REV | ILLIQ | TSKEW | SSKEW | ISKEW | *E*(TSKEW) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 0.1330 | | | |
| | | | | | | | (2.56) | | | |
| | | | | | | | | 0.2436 | | |
| | | | | | | | | (0.84) | | |
| | | | | | | | | | 0.1324 | |
| | | | | | | | | | (2.53) | |
| −0.0551 | 0.0474 | −0.1292 | 0.3194 | 0.6983 | −0.0712 | 0.0288 | 0.0436 | | | |
| (−5.36) | (1.90) | (−2.89) | (4.54) | (4.95) | (−13.29) | (3.53) | (1.67) | | | |
| −0.0538 | 0.0473 | −0.1318 | 0.3214 | 0.7057 | −0.0711 | 0.0290 | | 0.5202 | | |
| (−5.17) | (1.89) | (−2.89) | (4.53) | (4.98) | (−13.28) | (3.54) | | (0.91) | | |
| −0.0551 | 0.0475 | −0.1294 | 0.3195 | 0.6981 | −0.0712 | 0.0287 | | | 0.0426 | |
| (−5.36) | (1.90) | (−2.90) | (4.54) | (4.95) | (−13.29) | (3.53) | | | (1.63) | |
| | | | | | | | | | | 1.5188 |
| | | | | | | | | | | (3.92) |
| −0.0524 | 0.0435 | −0.0112 | 0.3366 | 0.5073 | −0.0752 | 0.0108 | | | | 0.9947 |
| (−4.82) | (1.30) | (−0.10) | (3.55) | (2.68) | (−12.06) | (0.05) | | | | (0.87) |

approximately −0.055, slightly smaller in magnitude than the −0.064 reported in Table 7, but still economically very significant and statistically significant at all conventional levels, with *t*-statistics above 5.0 in magnitude. In all the specifications, the coefficients on the skewness variables are positive, the opposite of the sign one would expect if investors have a preference for positive skewness. However, in the full specifications these average coefficients are statistically insignificant. The results for systematic skewness (co-skewness) differ from the significant negative relation found in Harvey and Siddique (2000) and Smith (2007), presumably due to differences in the methodology. For idiosyncratic skewness, we cannot replicate the negative and significant relation found in Zhang (2005) and Boyer, Mitton, and Vorkink (2010). Again differences in methodology presumably account for the discrepancy, a key difference being that both papers predict only portfolio returns, not the returns on individual securities.

For our purposes, however, the message of Tables 11 and 13 is clear. There is no evidence that the effect of extreme positive returns that we show is subsumed by available measures of skewness.

## 5. Conclusion

We find a statistically and economically significant relation between lagged extreme positive returns, as measured by the maximum daily return over the prior month or the average of the highest daily returns within the month, and future returns. This result is robust to controls for numerous other potential risk factors and control variables. Of particular interest, inclusion of our MAX variable reverses the anomalous negative relation between idiosyncratic volatility and returns in Ang, Hodrick, Xing, and Zhang (2006, 2009). We interpret our results in the context of a market with poorly diversified yet risk-averse investors who have a preference for lottery-like assets. In fact, it may be the preference for lottery-like payoffs that causes underdiversification in the first place, since well-diversified equity portfolios do not exhibit this feature. Thus, the expected returns on stocks that exhibit extreme positive returns are low but, controlling for this effect, the expected returns on stocks with high idiosyncratic risk are high.

Why is the effect we report not traded away by other well-diversified investors? Exploiting this phenomenon would require shorting stocks with extreme positive returns. The inability and/or unwillingness of many investors to engage in short-selling has been discussed extensively in the literature. Moreover, stocks with extreme positive returns are small and illiquid, on average, suggesting that transactions costs may be a serious impediment to implementing the relevant trading strategy. Finally, these small stocks tend to be held and traded by individual investors, rather than by institutions who might attempt to exploit this phenomenon.

We also present some evidence that stocks with extreme negative returns exhibit the reverse effect, i.e., investors find them undesirable and hence, they offer higher future returns. However, this phenomenon is not robust in all our cross-sectional regression specifications, and it appears to be concentrated in a smaller subsample of stocks than the effect of extreme positive returns. Of course, since exploiting this anomaly does not require taking a short position, one might expect the effect to be smaller than for stocks with extreme positive returns due to the actions of well-diversified traders.

While the extreme daily returns we exploit are clearly idiosyncratic, we make no effort to classify them further. In other words, we do not discriminate between returns due to earnings announcements, takeovers, other corporate events, or releases of analyst recommendations. Nor do we distinguish price moves that occur in the absence of any new public information. Interestingly, the preponderance of existing evidence indicates that stocks under-react rather than over-react to firm-specific news[14]; therefore, if the extreme positive returns were caused by good news, one should expect to see the reverse of the effect that we show. Given the magnitude and robustness of our results, this presents a potentially fruitful avenue of further research. Investigating the time-series patterns in the return premiums we document is also of interest. For example, it is conceivable that the magnitude of these premiums is affected by investor sentiment (Baker and Wurgler, 2007).

## Appendix. Variable definitions

MAXIMUM: MAX is the maximum daily return within a month:

$$\mathrm{MAX}_{i,t} = \max(R_{i,d}), \quad d = 1,\ldots,D_t, \tag{2}$$

where $R_{i,d}$ is the return on stock $i$ on day $d$ and $D_t$ is the number of trading days in month $t$.

MINIMUM: MIN is the negative of the minimum daily return within a month:

$$\mathrm{MIN}_{i,t} = -\min(R_{i,d}), \quad d = 1,\ldots,D_t, \tag{3}$$

where $R_{i,d}$ is the return on stock $i$ on day $d$ and $D_t$ is the number of trading days in month $t$.

TOTAL VOLATILITY: The total volatility of stock $i$ in month $t$ is defined as the standard deviation of daily returns within month $t$:

$$\mathrm{TVOL}_{i,t} = \sqrt{\mathrm{var}(R_{i,d})}. \tag{4}$$

BETA: To take into account nonsynchronous trading, we follow Scholes and Williams (1977) and Dimson (1979) and use the lag and lead of the market portfolio as well as the current market when estimating beta:

$$R_{i,d}-r_{f,d} = \alpha_i + \beta_{1,i}(R_{m,d-1}-r_{f,d-1}) + \beta_{2,i}(R_{m,d}-r_{f,d}) + \beta_{3,i}(R_{m,d+1}-r_{f,d+1}) + \varepsilon_{i,d}, \tag{5}$$

where $R_{i,d}$ is the return on stock $i$ on day $d$, $R_{m,d}$ is the market return on day $d$, and $r_{f,d}$ is the risk-free rate on day $d$.[15] We estimate Eq. (5) for each stock using daily returns within a month. The market beta of stock $i$ in month $t$ is defined as $\hat{\beta}_i = \hat{\beta}_{1,i} + \hat{\beta}_{2,i} + \hat{\beta}_{3,i}$.

IDIOSYNCRATIC VOLATILITY: To estimate the monthly idiosyncratic volatility of an individual stock, we assume a single-factor return-generating process:

$$R_{i,d}-r_{f,d} = \alpha_i + \beta_i(R_{m,d}-r_{f,d}) + \varepsilon_{i,d}, \tag{6}$$

where $\varepsilon_{i,d}$ is the idiosyncratic return on day $d$. The idiosyncratic volatility of stock $i$ in month $t$ is defined as the standard deviation of daily residuals in month $t$:

$$\mathrm{IVOL}_{i,t} = \sqrt{\mathrm{var}(\varepsilon_{i,d})}. \tag{7}$$

SIZE: Following the existing literature, firm size is measured by the natural logarithm of the market value of equity (a stock's price times shares outstanding in millions of dollars) at the end of month $t-1$ for each stock.

BOOK-TO-MARKET: Following Fama and French (1992), we compute a firm's book-to-market ratio in month $t$ using the market value of its equity at the end of December of the previous year and the book value of common equity plus balance-sheet deferred taxes for the firm's latest fiscal year ending in the prior calendar year.[16]

INTERMEDIATE-TERM MOMENTUM: Following Jegadeesh and Titman (1993), the momentum variable for each stock in month $t$ is defined as the cumulative return on the stock over the previous 11 months starting two months ago, i.e., the cumulative return from month $t-12$ to month $t-2$.

SHORT-TERM REVERSAL: Following Jegadeesh (1990) and Lehmann (1990), the reversal variable for each stock in month $t$ is defined as the return on the stock over the previous month, i.e., the return in month $t-1$.

ILLIQUIDITY: Following Amihud (2002), we measure stock illiquidity for each stock in month $t$ as the ratio of the absolute monthly stock return to its dollar trading volume:

$$\mathrm{ILLIQ}_{i,t} = |R_{i,t}|/\mathrm{VOLD}_{i,t}, \tag{8}$$

where $R_{i,t}$ is the return on stock $i$ in month $t$, and $VOLD_{i,t}$ is the respective monthly trading volume in dollars.

[14] See Daniel, Hirshleifer, and Subrahmanyam (1998) for a survey of some of this literature.

[15] In our empirical analysis, $R_{m,d}$ is measured by the CRSP daily value-weighted index and $r_{f,d}$ is the one-month T-bill return available at Kenneth French's online data library.

[16] To avoid issues with extreme observations, following Fama and French (1992), the book-to-market ratios are winsorized at the 0.5% and 99.5% levels, i.e., the smallest and largest 0.5% of the observations on the book-to-market ratio are set equal to the 0.5th and 99.5th percentiles, respectively.

TOTAL SKEWNESS: The total skewness of stock $i$ for month $t$ is computed using daily returns within year $t$:

$$\mathrm{TSKEW}_{i,t} = \frac{1}{D_t}\sum_{d=1}^{D_t}\left(\frac{R_{i,d}-\mu_i}{\sigma_i}\right)^3, \tag{9}$$

where $D_t$ is the number of trading days in year $t$, $R_{i,d}$ is the return on stock $i$ on day $d$, $\mu_i$ is the mean of returns of stock $i$ in year $t$, and $\sigma_i$ is the standard deviation of returns of stock $i$ in year $t$.

SYSTEMATIC and IDIOSYNCRATIC SKEWNESS: Following Harvey and Siddique (2000), we decompose total skewness into idiosyncratic and systematic components by estimating the following regression for each stock:

$$R_{i,d}-r_{f,d} = \alpha_i + \beta_i(R_{m,d}-r_{f,d}) + \gamma_i(R_{m,d}-r_{f,d})^2 + \varepsilon_{i,d}, \tag{10}$$

where $R_{i,d}$ is the return on stock $i$ on day $d$, $R_{m,d}$ is the market return on day $d$, $r_{f,d}$ is the risk-free rate on day $d$, and $\varepsilon_{i,d}$ is the idiosyncratic return on day $d$. The idiosyncratic skewness (ISKEW) of stock $i$ in year $t$ is defined as the skewness of daily residuals $\varepsilon_{i,d}$ in year $t$. The systematic skewness (SSKEW) or co-skewness of stock $i$ in year $t$ is the estimated slope coefficient $\hat{\gamma}_{i,t}$ in Eq. (10).

## References


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

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. Journal of Finance 61, 259–299.

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence. Journal of Financial Economics 91, 1–23.

Arditti, F.D., 1967. Risk and the required return on equity. Journal of Finance 22, 19–36.

Arditti, F.D., 1971. Another look at mutual fund performance. Journal of Financial and Quantitative Analysis 6, 909–912.

Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. Journal of Economic Perspectives 21, 129–151.

Bali, T.G., Cakici, N., 2008. Idiosyncratic volatility and the cross-section of expected returns. Journal of Financial and Quantitative Analysis 43, 29–58.

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

Barone-Adesi, G., 1985. Arbitrage equilibrium with skewed asset returns. Journal of Financial and Quantitative Analysis 20, 299–313.

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

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

Calvet, L.E., Campbell, J.Y., Sodini, P., 2007. Down or out: assessing the welfare costs of household investment mistakes. Journal of Political Economy 115, 707–747.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Conrad, J., Dittmar, R.F., Ghysels, E., 2008. Skewness and the bubble. Unpublished working paper, University of North Carolina at Chapel Hill.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839–1885.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross-sectional variation in stock returns. Journal of Finance 52, 1–33.

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

Fama, E.F., French, K.R., 1992. Cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E.F., French, K.R., 2008. Average returns, B/M, and share issues. Journal of Finance 63, 2971–2995.

Fama, E.F., MacBeth, J.D., 1973. Risk and return: some empirical tests. Journal of Political Economy 81, 607–636.

Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. Journal of Finance 64, 2023–2052.

Friend, I., Westerfield, R., 1980. Co-skewness and capital asset pricing. Journal of Finance 35, 897–913.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. Journal of Financial Economics 91, 24–37.

Garrett, T.A., Sobel, R.S., 1999. Gamblers favor skewness, not risk: further evidence from United States' lottery games. Economics Letters 63, 85–90.

Goetzmann, W.N., Kumar, A., 2008. Equity portfolio diversification. Review of Finance 12, 433–463.

Han, B., Kumar, A., 2008. Retail clienteles and the idiosyncratic volatility puzzle. Unpublished working Paper, University of Texas at Austin.

Harvey, C., Siddique, A., 1999. Autoregressive conditional skewness. Journal of Financial and Quantitative Analysis 34, 465–487.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263–1295.

Huang, W., Liu, Q., Rhee, G., Zhang, L., 2010. Return reversals, idiosyncratic risk, and expected returns. Review of Financial Studies 23, 147–168.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Kane, A., 1982. Skewness preference and portfolio choice. Journal of Financial and Quantitative Analysis 17, 15–25.

Keim, D.B., 1999. An analysis of mutual fund design: the case of investing in small-cap stocks. Journal of Financial Economics 51, 173–194.

Kraus, A., Litzenberger, R.H., 1976. Skewness preference and the valuation of risk assets. Journal of Finance 31, 1085–1100.

Kumar, A., 2009. Who gambles in the stock market? Journal of Finance 64, 1889–1933

Lehmann, B., 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics 105, 1–28.

Lim, K.-G., 1989. A new test of the three-moment capital asset pricing model. Journal of Financial and Quantitative Analysis 24, 205–216.

Lintner, J., 1965. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13–37.

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

Mossin, J., 1966. Equilibrium in a capital asset market. Econometrica 34, 768–783.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Nyberg, P., 2008. The dynamic behavior of the idiosyncratic volatility discount: aggregate idiosyncratic volatility and return reversals revisited. Unpublished working paper, Hanken School of Economics.

Odean, T., 1999. Do investors trade too much? American Economic Review 89, 1279–1298

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Patel, N.R., Subrahmanyam, M.G., 1978. Utility theory and participation in unfair lotteries. Journal of Economic Theory 19, 555–557.

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data. Journal of Financial Economics 5, 309–327.

Sears, R.S., Wei, J.K.C., 1985. Asset pricing, higher moments, and the market risk premium: a note. Journal of Finance 40, 1251–1253.

Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19, 425–442.

Smith, D.R., 2007. Conditional coskewness and asset pricing. Journal of Empirical Finance 14, 91–119.

Spiegel, M.I., Wang, X., 2005. Cross-sectional variation in stock returns: liquidity and idiosyncratic risk. Unpublished working paper, Yale University.

Thaler, R.H., Ziemba, W.T., 1988. Parimutuel betting markets: racetracks and lotteries. Journal of Economic Perspectives 2, 161–174.

Tversky, A., Kahneman, D., 1992. Advance in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297–323.

Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. Review of Economic Studies 77, 779–805.

Walker, I., Young, J., 2001. An economist's guide to lottery design. Economic Journal 111, F700–F722.

Zhang, Y., 2005. Individual skewness and the cross-section of expected returns. Unpublished working paper, Yale University.