Gambling Preferences for Loser Stocks*

Peixuan Yuan

Rutgers Business School

Rutgers University†

This version: June, 2020

Abstract

I discover that investors' preferences for gambling mainly involve stocks that have performed poorly in the past three months, as lottery-like stocks with poor performance are much more likely to generate large payoffs than those with good performance (61.53% vs. 40.17%). Furthermore, lotto investors tend to believe that lottery-like stocks with poor performance may have a vigorous rebound shortly, while those with good performance may be less likely to produce a highly positive return given their high prices. Therefore, lottery-like stocks with poor performance have a highly effective lottery-like look, and thus they attract lotto investors. On the other hand, loser stocks without lottery-like features may continue to perform poorly. Overly optimistic (pessimistic) beliefs about stocks with (without) lottery-like features result in a pronounced lottery premium among loser stocks.

Keywords: Gambling preference; Past performance; Effective lottery-like features; Lotto investors

JEL classification: G10, G11, G12, G14

---

*I thank Azi Ben-Rephael, Ren-raw Chen, Priyank Gandhi, Zhengzi (Sophia) Li, Yangru Wu, Ken Zhong and seminar participants at Rutgers Business School for helpful conversations and comments. All remaining errors are my own.

†Piscataway, NJ, 08854. Phone: (609)997-1398. Email: peixuan.yuan@rutgers.edu

1. Introduction

Recent research has established the important role of investors' gambling preferences in determining stock returns, i.e., lottery-like stocks that offer a small chance of a huge payoff tend to earn low returns.[1] Motivated by the inference of the Merton (1974) capital structure model that loser firms, which are often extremely stressed and at risk of bankruptcy, are effectively an out-of-the-money option on the underlying firm value, in this paper, I further investigate whether a stock's past performance has a significant impact on its lottery premium.[2]

Prior studies utilized stocks' past characteristics to represent their future likelihood of having huge payoffs. For example, Bali et al. (2011) used maximum daily return (MAX) in the previous month as a proxy for lottery-like features, based on the assumption that stocks with extremely high returns in the past are likely to have huge payoffs in the future. However, in this paper, I show that such an assumption depends greatly on the stocks' past performance. Specifically, I find that high-MAX stocks that performed poorly in the past three months have a 61.53% probability of generating a huge payoff in the next month, whereas high-MAX stocks that performed well only have a probability of 40.17% of having an extremely high return.[3] The difference between these two probabilities is 21.36 percentage points, with a t statistic of 29.68.

Furthermore, lotto investors may prefer to buy high-MAX stocks with poor performance than those with excellent performance, as they believe stocks with poor performance have more space for upwards movement. Thus, lotto investors are less likely to buy a winner stock with a maximum daily return of 25% in the last month than a loser stock with a maximum daily return of 20%, although the winner stock has a higher MAX, they might think the current price of the winner stock is too high for another substantial positive jump. They prefer the loser stock, as its current

[1] Kumar (2009) shows that individual investors are attracted to lottery-type stocks with low prices, high idiosyncratic volatility, and high idiosyncratic skewness. Bali, Cakici, and Whitelaw (2011) find that high-MAX stocks underperform low-MAX stocks.

[2] In the sense of Merton's (1974) capital structure model, a share of common stock is a call option on the underlying firm's assets when there is debt in the capital structure. The underlying values among past losers have generally suffered severely and are, therefore, potentially much closer to a level at which the option convexity is strong. Thus, loser firms which are often extremely stressed and at risk of bankruptcy are effectively an out-of-the-money option on the underlying firm value. Past winners, in contrast, are likely in the money. Daniel and Moskowitz (2016) also find that momentum strategies can experience infrequent and persistent strings of negative returns because of the option-like payoffs of past losers.

[3] First, I sort stocks into quintile portfolios by MAX in month t. Next, I form 25 portfolios in month $t-1$ by independently sorting stocks based on MAX and stock past performance (cumulative return of the past three months). Finally, I calculate the proportion of high-MAX stocks with a good (poor) performance that belong to the high-MAX portfolio for month t.

price is very low and, thus, it is more likely to jump back to its previous level. In other words, lotto investors tend to believe that lottery-like stocks with poor performance are relatively cheaper and may enjoy vigorous rebounds shortly. Therefore, high-MAX stocks with poor performance have strong effective lottery-like features and thus attract lotto investors. On the other hand, low-MAX stocks with poor performance may continue to perform poorly. Overly optimistic (pessimistic) beliefs about stocks with (without) lottery-like features result in pronounced lottery premiums among loser stocks.

To test the conjecture that lottery premium is stronger among loser stocks, I first perform conditional bivariate portfolio sorts. Specifically, I perform a sequential sort by creating three portfolios (30%-40%-30%) ranked on past stock performance (PFM) measured by the stocks' cumulative returns across months t-3 and t-1.[4] The bottom 30% stocks are denoted as losers, and the top 30% stocks as winners. Then, within each PFM portfolio, I form a second set of quintile portfolios ranked on the stock's lottery-like features in month t-1, and I hold each portfolio in month t. I use two measures of the lottery-like feature. First, following Bali et al. (2011), I use MAX as a proxy for a stock's ex-ante lottery-like features. Second, following Kumar (2009) and Bali, Hirshleifer, Peng, and Tang (2019), I use an index of lottery likeness (LTRY) that defines lottery stocks as low-priced stocks with high idiosyncratic volatility (IVOL) and high idiosyncratic skewness (ISKEW). I find that the lottery premium is indeed very high among stocks with low past returns (losers), but very low among stocks with high past returns (winners). Specifically, the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS)[5] alpha spread is -2.13% per month with a t statistic of -7.73 for loser stocks, whereas the alpha spread is -0.67% (t statistic = -2.70) for winner stocks when I use MAX as the proxy for lottery-like features.

Furthermore, I observe that, in the loser group, low-MAX stocks have a significant positive alpha of 0.57% per month (t = 5.22), and high-MAX stocks have a significant negative alpha of -1.56% per month (t = 6.64). Therefore, the lottery premium is driven by both the underperformance of high-MAX stocks and the outperformance of low-MAX stocks. Similar results are obtained when lottery-like payoffs are proxied by LTRY. I then investigate the characteristics of loser stocks, which have relative smaller firm sizes, lower return reversal, lower price per share, higher illiquidity

[4]The results are robust to alternative measurements of past stock performance.

[5]In this paper, I apply the Fama-French-Carhart-Pastor-Stambaugh model to calculate the risk-adjusted return. My results are robust to alternative risk factor models.

(ILLIQ), higher IVOL, and lower capital gain overhang (CGO). To test whether my findings are due to these characteristics, I employ trivariate portfolio analysis. I find that the results from the bivariate sorts cannot be explained by any one of these characteristics.

To find out why the lottery premium is significant among loser stocks, I first examine the relationship between PFM and future extreme returns. I define the monthly price boost (crash) measure as an indicator variable that equals one for a firm-month that experiences one or more price boost (crash) days during that month and zero otherwise. I then perform a cross-sectional regression by first conducting a monthly logistic regression of the one-month-ahead price boost or crash measure on current PFM, controlling for a set of firm characteristics. The results from the regressions indicate that the loser stocks have a high probability of future extremely high returns, but not necessarily extremely low returns. Such attributes of loser stocks may attract investors with gambling preferences.

I further investigate the proportion of realized extremely high returns among loser stocks by calculating the portfolio transition matrix. I observe that the loser stocks in month t-1 have a high (low) probability of landing in the high-MAX (low-MAX) portfolio in month t. However, there is no significant pattern with winner stocks. Thus, investors seeking huge short-term payoffs might have more interest in loser stocks.

Kumar (2009), Han and Kumar (2013), and Kumar, Page, and Spalt (2016) provide evidence that investors with a stronger propensity to gamble tend to be more attracted to lottery stocks. This suggests that the lottery premium is greater among individual investors than institutional investors (see, e.g., Bali et al. (2019) and Lin and Liu (2017)). Therefore, this attraction to retail investors is likely to push up the stock prices and, therefore, induce overvaluation. I next examine whether retail holdings are higher in loser stocks and especially in loser stocks with lottery-like features. I sort all stocks in the sample into quintiles based on PFM and calculate the average level of retail holdings in month t-1 (RHLD). To mitigate the deposition effect, I also calculate the changes in RHLD over the months t-3 to t-1 (ΔRHLD). I observe that RHLD is largest in loser stocks, and, more importantly, that ΔRHLD is significantly positive (0.57%) in Quintile 1 (Loser) and monotonically decreases to -1.28% in Quintile 5 (Winner). Furthermore, I conduct conditional bivariate sorts based on PFM and MAX, and I find that RHLD is the largest in the portfolio with the lowest PFM and highest MAX. What is more, loser stocks with lottery-like features experienced a significant increase in retail holdings during the previous quarter. Therefore, retail investors are indeed attracted by loser stocks, especially those with lottery-like features.

The biased beliefs framework proposed in Engelberg, Mclean, and Pontiff (2018) suggests that investors may be overly optimistic (pessimistic) about certain groups of firms. The arrival of news may then force them to update their biased beliefs rapidly, resulting in abnormal returns over a short event window during which expected returns are close to zero. To test my conjecture that lotto investors' biased beliefs induce a pronounced lottery premium among loser stocks, I study the market reaction to firm-level earnings announcements. I conduct conditional bivariate sorts as before and I calculate value-weighted portfolio-abnormal returns around earnings announcement date in the following month.[6] I find that the three-day abnormal return spread between high-MAX and low-MAX stocks in loser group is economically and statistically significant (-1.16% with a t statistic of -3.92), while the spread in the winner group is not. Thus, investors' biased beliefs appear to apply only to loser stocks. Furthermore, investors are overly optimistic (pessimistic) about stocks with high (low) maximum daily returns, resulting in a highly significant negative (positive) abnormal three-day event-time return.

Our findings show that the significant lottery premium in loser stocks is driven by both the underperformance of high-MAX stocks and the outperformance of low-MAX stocks, and that high-MAX stocks have the largest effect. I note that the short-sale constraint is the necessary condition for optimistic belief to induce overpricing. It is then natural to expect an optimistic belief of gambling investors to induce overpricing only among stocks with binding short-sale constraints. I examine this hypothesis by conducting similar conditional bivariate sorts analysis, but with stocks with options. I find that the alpha of high-MAX stocks in the loser group is no more significant, with a value of -0.3% and a t statistic of -0.89. However, the alpha of low-MAX stocks in the loser group remains significant, with a value of 0.76 and a t statistic of 3.20. Therefore, optimistic beliefs about high-MAX loser stocks, combined with short-sale constraints, drive the short side of the lottery premium, whereas pessimistic beliefs about low-MAX loser stocks contribute to the long side of the lottery premium.

Baker and Wurgler (2004) find that investor sentiment explains the cross-section of stock returns,

[6]Abnormal announcement returns are defined as the difference between the actual buy-and-hold return over the event days $t = -1$ to $t = 1$ and the expected Fama and MacBeth (1973) three-factor buy-and-hold return.

with high sentiment being a significant predictor of the returns of more speculative stocks such as those of small firms, new firms, and highly volatile firms. When sentiment is high, subsequent returns on these stocks tend to be low and vice versa. Stambaugh, Yu, and Yuan (2012) document that investor bias and mispricing is stronger during periods of high investor sentiment. Therefore, if investor-biased belief manifests in loser stocks, we should observe that the lottery premium is notably stronger during periods of high investor sentiment. In other words, the difference in lottery premium between loser and winner stocks should be much higher when investors have more propensity to gamble (high sentiment periods). Furthermore, lotto investors are more likely to overvalue loser stocks with lottery-like features, while disliking those without lottery-like features. Consistent with prior findings, I observe that when sentiment is high, investors are more optimistic (pessimistic) about the future payoffs of loser stocks with (without) lottery-like features than when sentiment is low.

I also perform Fama and MacBeth (1973) regression tests to control for various well-known return predictors. After controlling for the market beta (BETA), firm size, book-to-market ratio (BM), price momentum (MOM), return reversals (REV), ILLIQ, IVOL, realized skewness (RSKEW), realized kurtosis (RKURT), co-skewness (CSK), and co-kurtosis (CKT), the estimated coefficients on the interaction of lottery-like features and loser stock indicators are all consistent with the results from portfolio sorting analysis. Specifically, the average coefficient of the interaction between MAX and loser stocks indicators from the monthly regressions of Model 1 in Table 12 is -18.62 with a t statistic of -7.13, implying that a one standard deviation increase in MAX in the previous month corresponds to a 4.25% decrease in monthly returns for loser stocks. The coefficient of MAX, however, becomes insignificant when I include the interaction term.

This paper is most closely related to An, Wang, Wang, and Yu (forthcoming). An et al. argue that reference-dependent preferences cause skewness-related anomalies to be concentrated among stocks where investors have lost money, as investors who are in losses are less likely to sell lottery stocks to make up their losses. My research, however, is distinct from theirs in the following aspects: (1) They argue that the reason for overvaluation of the lottery assets is that the current holders facing capital losses are unwilling to sell. However, under my framework, I find that lottery assets among loser stocks are overpriced, as their retail holdings have experienced a significant increase. Thus, lotto investors are actively buying lottery-like stocks with poor performance. (2) Their argument cannot explain my finding that loser stocks without lottery-like features have a significant positive alpha. However, I show that investors' pessimistic beliefs induce the outperformance of such stocks. Most importantly, my results are robust after controlling for capital gain overhang (CGO).

The paper proceeds as follows. Section 2 discusses the data and the definition of variables. Section 3 presents preliminary results from sorting the stocks into quintiles based on lottery-like features. Section 4 discusses the empirical findings and their possible explanation. Section 5 extends the portfolio-level analysis by performing Fama and MacBeth (1973) cross-sectional regressions. Section 6 examines the robustness of my results. Section 7 concludes.

2. Data and Methodology

I obtain the daily stock prices and returns data from the Center for Research in Security Prices (CRSP). I select all stocks traded on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the National Association of Securities Dealers Automated Quotations (NASDAQ) that are classified as ordinary common shares (CRSP codes 10 and 11) and I exclude closed-end funds and real estate investment trusts (Standard Industrial Classification codes 6720-6730 and 6798) for the sample period between July 1965 and December 2017. Following Amihud (2002) and Zhang (2006), I exclude penny stocks with prices below $1. Additionally, I required a market capitalization of at least $225 million (Baltussen, van Bekkum, and van der Grient (2018)). These two thresholds serve to eliminate the most illiquid stocks that exhibit potential microstructure problems and that may bias the results. I adjust for delisting returns following Shumway (1997) and Shumway and Warther (1999). I obtain accounting variables from the merged CRSP-Compustat database. The institutional ownership data come from Thompson 13F filings from 1980 through 2017. The excess market returns (MKT) and the size, book to market, and momentum factors, namely, small-minus-big (SMB), high-minus-low (HML), and momentum winner-minus-loser (UMD) come from Kenneth French's data library. The liquidity factor (PS) comes from Lubos Pastor's data library.

2.1. *Definitions of Key Variables*

Following Bali et al. (2019), I construct two types of proxies for lottery-like features. The primary proxy for the stock's ex-ante lottery-like feature is the maximum daily return of the stock in the previous month, denoted MAX (Bali et al. (2011)). My second proxy for lottery-like payoffs is adapted from Kumar (2009). For each month, I sort stocks into 50 bins by price per share (PRC) in descending order such that stocks in the lowest bin (i.e., a PRC portfolio rank of 50) have the lowest price per share. I also independently sort stocks into 50 bins by IVOL and ISKEW in ascending order. IVOL and ISKEW are, respectively, the standard deviation and the skewness of residuals from the time-series regression of daily stock returns against the market (MKT), size (SMB), and book-to-market (HML) factors in a month. I then construct a lottery index, denoted LTRY, by summing up the ranks of the PRC, IVOL, and ISKEW portfolios. The lottery index thus has an integer value in the range of 3 to 150, and it increases with a stock's lottery feature.

I measure PFM using stock's cumulative returns from months t-3 to t-1. Furthermore, I define RHLD as one minus the quarterly fractional institutional ownership and I denote a change in RHLD over the quarter as ΔRHLD. Following Cremers and Nair (2005), the quarterly institutional ownership is set to zero if it is missing from the database.

2.2. *Control variables*

I use a number of well-known cross-sectional return predictors as control variables in Fama and MacBeth (1973) regressions. Specifically, following Fama and MacBeth (1973), I estimate stock i's BETA using its monthly returns over the prior 60 months if available (or a minimum of 24 months), and I compute the stock's size (SIZE) as the product of the price per share and the number of shares outstanding. The BM at the end of June of year t is computed as the book value of stockholders' equity, plus deferred taxes and investment tax credit (if available), minus the book value of the preferred stock for the last fiscal year ending in t-1, scaled by the market value of equity at end of December of year $t-1$. Following Jegadeesh and Titman (1993), MOM is the cumulative return of a stock over 11 months ending a month before portfolio formation.

Following Harvey and Siddique (2000), the stock's monthly CSK is defined as:

$$\text{CSK}_{i,t} = \frac{\frac{1}{D}\sum_d (R_{i,d} - \mu_i)(\text{MKT}_d - \mu_{\text{MKT}})^2}{\sqrt{\text{var}(R_{i,d})}\text{var}(\text{MKT}_d)} \tag{1}$$

where $R_{i,d}$ is the return of stock i on day d, μ_i is the average excess return on the stock, and D denotes the number of trading days in the examination period.

Similarly, the stock's monthly CKT is defined as:

$$\text{CKT}_{i,t} = \frac{\frac{1}{D}\sum_d (R_{i,d} - \mu_i)(\text{MKT}_d - \mu_{\text{MKT}})^3}{\sqrt{\text{var}(R_{i,d})}\text{var}(\text{MKT}_d)^{3/2}} \tag{2}$$

Following Amihud (2002), I measure the ILLIQ of stock i in month t, denoted ILLIQ, as the average daily ratio of the absolute stock return to the dollar trading volume within the month:

$$\text{ILLIQ}_{i,t} = \frac{1}{D}\sum_{d=1}^{D} \frac{|R_{i,d}|}{\text{volume}_{i,d} \times \text{price}_{i,d}} \tag{3}$$

where $R_{i,d}$ is the return of stock i on day d, $\text{volume}_{i,d}$ is the daily trading volume, and $\text{price}_{i,d}$ is the daily price. I then transform the ILLIQ by its natural log to reduce skewness.

Following Ang, Hodrick, Xing, and Zhang (2006), the monthly IVOL of stock i is computed as the standard deviation of the residuals from the regression based on the daily return regression:

$$R_{i,d} - r_{f,d} = \alpha_i + \beta_i^{\text{MKT}}\text{MKT}_d + \beta_i^{\text{SMB}}\text{SMB}_d + \beta_i^{\text{HML}}\text{HML}_d + \epsilon_{i,d} \tag{4}$$

where $R_{i,d}$ is the return of stock i on day d, $r_{f,d}$ is the risk-free rate on day d. MKT, SMB and HML are, respectively, the excess return on the market portfolio and the return on two long/short portfolios that capture size and book-to-market effects (Fama and French (1993)).

2.3. Methodology

Our primary methodology involves sorting into quintiles based on ascending order of lottery-like features (MAX or LTRY) in month t-1 and evaluating returns in month t. I form a zero-investment portfolio by taking long positions on the bottom quintile of stocks in the past month and shorting the stocks in the top quintile. To examine the relation between PFM and investors' gambling preferences, I perform a sequential sort by creating three portfolios (30%-40%-30%) ranked on PFM. The bottom 30% are denoted as losers, and the top 30% as winners. Then, within each PFM portfolio, I form a second set of quintile portfolios ranked on lottery-like features (MAX or LTRY). This creates a set of portfolios with similar past performance and spreads in stocks' lottery-like features, and thus I can examine expected return differences due to investors' gambling preference rankings in stocks with different past performance. I hold these portfolios for one month (see Figure 1 for the timeline).

3. Preliminary Analysis

I begin my analysis by replicating prior findings on investors' gambling preferences. More specifically, I sort stocks based on their lottery-like features (MAX or LTRY) over the past month to form quintile portfolios. Quantile 1 has the lowest MAX or LTRY in the past month, and Quintile 5 has the highest MAX or LTRY in the past month. Table 1 presents the value-weighted results from July 1965 to December 2017. The first two columns report the results of portfolios sorted by MAX, and the next two columns report the results of portfolios sorted by LTRY. The first and third columns report the value-weighted average monthly excess return for each quintile and the value-weighted average return difference between Quintile 5 (high) and Quintile 1 (low). The second and fourth columns present the corresponding risk-adjusted returns (alphas) relative to the FFCPS, which includes the Fama-French three-factor buy-and-hold return, UMD, and PS. Newey and West (1987) t statistics are in parentheses.

The first two columns confirm the results of Bali et al. (2011) that the high-MAX portfolio (Quintile 5) has lower post-formation returns than the low-MAX portfolio (Quintile 1). The mean monthly excess returns of these extreme portfolios are 0.12% and 0.60% respectively, giving a return spread of -0.47% (-1.71 t statistics). Adjusting for risk using the FFCPS model magnifies the MAX effect to an alpha spread of -0.72%, which is significant with a t statistic of -3.15. Therefore, Quintile 5 stocks generate 70 basis points per month lower risk-adjusted returns than stocks in Quintile 1. Furthermore, it appears that the negative alpha spread between high-MAX and low-MAX stocks is driven by the underperformance of high-MAX stocks: the FFCPS alpha of low-MAX stocks is not significant, whereas the alpha of high-MAX stocks is negative, economically large, and statistically significant, at -59 basis points per month (-2.78 t statistic). The value-weighted portfolio analyses are similar for LTRY and MAX. Overall, my preliminary analysis confirms the prior findings on investors' gambling preferences, which cause lottery stocks to be overpriced and which subsequently generate lower returns.

4. Stocks' Past Performance and Lottery Premium

In this section, I provide my main findings on investor's gambling preference. Specifically, I apply the methodology described in Section 2.3 to examine the lottery effect when conditioned on stock's past performance. Basically, for each month, I group all stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending sort of stock's past returns from month t-3 to month t-1 (PFM). The bottom 30% stocks are denoted as losers, and the top 30% stocks winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending sort of a lottery proxy (MAX or LTRY) calculated in month t-1. The intersections of the three PFM groups and the five lottery proxy groups generate 15 value-weighted portfolios. Finally, I calculate portfolios' mean holding period returns in month t.

Panel A of Table 2 presents the FFCPS alpha for each of the 15 portfolios. The first three columns report FFCPS alphas for loser, medium, and winner stocks, respectively. The last two rows present the alpha spreads between the high-MAX and low-MAX quintiles within each PFM group. I find that the lottery premium is very high among stocks with low past returns (losers), but very low among stocks with high past returns (winners). Specifically, the alpha spread is -2.13% per month with a t statistic of -7.73 for loser stocks, whereas the alpha spreads are -0.66% (-3.99 t statistic) and -0.67% (-2.70 t statistic) for medium and winner stocks, respectively.

I also notice that, in the loser group, stocks with low-MAX have a significant positive alpha of 0.57% per month ($t = 5.22$), and stocks with high-MAX have a significant negative alpha of -1.56% per month ($t = 6.64$). Therefore, the lottery premium is driven by both the underperformance of high-MAX stocks and the outperformance of low-MAX stocks, although most of the effect is from high-MAX stocks. This is contrary to the prior findings of univariate portfolio analysis that negative alpha spread between high-MAX and low-MAX stocks is driven only by the underperformance of high-MAX stocks.

The last column reports the alpha spread between loser and winner stocks. Low-MAX stocks with poor performance have a significantly higher alpha than those with good performance as the loser-minus-winner (LMW) alpha spread is 0.42% (t = 2.64). Furthermore, high-MAX stocks with poor performance have much lower alpha than those with good performance, since the LMW alpha spread is -1.04% (t = 3.88). As a result, the LMW alpha spread between the high-low portfolio for loser stocks and that for winner stocks is both economically and statistically significant (-1.46%, t = 5.11).

Similar results are obtained when lottery-like payoffs are proxied by LTRY. Panel B of Table 2 shows that in the loser group, the FFCPS alpha spread is -1.47% per month with a t statistic of -7.19 for LTRY-sorted portfolios. In contrast, in the winner group, the FFCPS spread is only -0.27% with a non-significant t statistic of -1.54. The LMW alpha spread between the high-low portfolio for loser stocks and that for winner stocks is also both economically and statistically significant (-1.20%, t = 5.31). Moreover, loser stocks with low-MAX have a significant positive alpha of 0.29% per month (t = 2.87), and those with high-MAX have a significantly negative alpha of -1.19% per month (t = 6.54). Overall, the results indicate that the lottery premium is most pronounced among loser stocks, and that the lottery premium in loser stocks is driven by both the underperformance of high-MAX stocks and the outperformance of low-MAX stocks.

4.1. Characteristics of loser and winner stocks

The significant difference in lottery premium between different stocks' PFM groups inspires us to explore the characteristics of different PFM stocks further. Table 3 presents the average characteristics of each PFM group. The first column shows that PFM is -0.16%, 0.01%, and 0.21% for loser, medium, and winner stocks, respectively. Other characteristics include BETA, SIZE, BM, REV, ILLIQ, MAX, IVOL, ISKEW, and PRC. I notice that BETA, BM, and ISKEW stay at similar levels for both loser and winner stocks. However, loser stocks have smaller SIZE, lower REV, lower PRC, higher ILLIQ, higher IVOL, and lower CGO. Therefore, I examine whether the characteristics of different PFM groups induce significant differences in lottery premium.

To control for different characteristics among PFM groups, I employ trivariate portfolio sort analysis. Specifically, I first perform a sequential sort by creating three portfolios with $30\% - 40\% - 30\%$ breakpoints ranked by control variables. Then, within each control variable portfolio, I group all stocks into three portfolios with $30\% - 40\% - 30\%$ breakpoints ranked by PFM. Next, I group PFM portfolios that have the rank cross the control portfolios into three aggregate PFM portfolios. As a result, the aggregate PFM portfolios have a similar level of control variables. Finally, within each aggregate portfolios, I form a set of quintile portfolios ranked on stock's lottery-like features.[7]

Panel A in Table 4 reports the results of trivariate portfolio sorts with MAX as the lottery proxy. As shown in the first row, I control for SIZE, PRC, REV, ILLIQ, IVOL, and CGO. For each control variable, I report FFCPS alphas for each portfolio in loser and winner stocks. The row denoted Average presents the mean value of the control variables in each PFM group. The first and second columns show that the average firm size is 11.68 for loser stocks and 11.74 for winner stocks. Therefore, loser stocks and winner stocks have a similar firm sizes. The last two rows present the alpha spread between the high-MAX and low-MAX quintiles within each PFM group. I continue to find that the lottery premium is very high among stocks with low past returns (losers), but very low among stocks with high past returns (winners). The FFCPS alpha spread in loser group varies from -1.53% (t = 5.52) when controlling for CGO to -2.04% (t = 7.24) when controlling for REV. All alpha spreads are statistically significant, with an average t statistic of -7.15. In contrast, the average of alpha spread in the winner group is -0.63% with a t statistic of -2.54.

Panel B in Table 4 presents a similar result when lottery-like payoffs are proxied by LTRY. The FFCPS alpha spread in loser group varies from -1.05% (t = 5.68) when controlling for CGO to -1.52% (t = 7.18) when controlling for REV. All alpha spreads are statistically significant with an average t statistic of -7.25. By contrast, the average alpha spread in the winner group is -0.11% with a t statistic of -0.64. Therefore, the lottery premium is pronounced in loser stocks, and no single characteristic explains this.

4.2. Loser stocks and future extreme returns

In the sense of the Merton (1974) capital structure model, a share of common stock is a call option on the underlying firm's assets when there is debt in the capital structure. The underlying

[7] The purpose of this exercise is to investigate whether the differences in characteristics among PFM groups can explain the results. Therefore, I do not control for the effect of different characteristics on lottery-like features. However, I still find the lottery premium is more significant among loser stocks when I control for both firm's characteristics and PFM. More specifically, I first perform a double sort by creating 9 (3×3) portfolios with $30\% - 40\% - 30\%$ breakpoints ranked by control variables and stock's past performance (PFM). Next, within each 9 portfolios, I form a set of quintile portfolios ranked on stock's lottery-like features (MAX, or LTRY), resulting in a total 45 ($3 \times 3 \times 5$) portfolios.

firm values among past losers have generally suffered severely and they are, therefore, potentially much closer to a level at which the option convexity is strong. Hence, the loser firms, which are often extremely stressed and at risk of bankruptcy, are effectively an out-of-the-money option on the underlying firm value. The past winners, in contrast, are likely in the money. Daniel and Moskowitz (2016) also find that momentum strategies can experience infrequent and persistent strings of negative returns because of the option-like payoffs of past losers. These arguments inspire us to examine the relation between PFM and the probability of having future extreme returns.

4.2.1. Cross-section regression analysis

I first define the monthly price boost (crash) measure as the indicator variable that equals one for a firm-month in which stocks experience one or more price boost (crash) days and zero otherwise. Like the definition of Hutton, Marcus, and Tehranian (2009) and Kim, Li, and Zhang (2011), price boost (crash) days in a given month are days in which the firm experiences daily returns that are 3.09 (0.1% for normal distribution) standard deviations above (below) the mean daily returns over the entire year.[8]

To examine the predictability of PFM for future price boost or crash probability, I perform a cross-sectional regression by first conducting a monthly logistic regression of the one-month-ahead price boost or crash measure on current PFM, controlling for a set of firm characteristics:

$$\mathrm{EXTREME}_{i,t} = \beta_{i,0} + \beta_{i,1}\mathrm{PFM}_{i,t-1} + \beta_{i,2}X_{i,t-1} + \epsilon_{i,t} \tag{6}$$

where $\mathrm{EXTREME}_{i,t}$ denotes the indicator of price boost or price crash for stock i in month t, $\mathrm{PFM}_{i,t-1}$ represents the stock i's cumulative return from month t-3 to month t-1, and $X_{i,t-1}$ is a vector of control variables for stock i in month t, including BETA, SIZE, REV, ILLIQ, IVOL, RSKEW, RKURT, CSK, and CKT. I then aggregate coefficients across all months and compute the Newey and West (1987) t statistics.

Panel A in Table 5 presents the results from cross-sectional regression where EXTREME denotes

[8] Our results are robust to firm-specific return, which is defined as the natural log of one plus the residual return from the regression:

$$R_{i,d} = \alpha_i + \beta_i^1\mathrm{MKT}_{d-2} + \beta_i^2\mathrm{MKT}_{d-1} + \beta_i^3\mathrm{MKT}_{d} + \beta_i^4\mathrm{MKT}_{d+1} + \beta_i^5\mathrm{MKT}_{d+2} + \epsilon_{i,d} \tag{5}$$

where $R_{i,d}$ is the return on stock i on day d and MKT is the market return.

the indicator of price boost. The first column shows that PFM has a significantly negative coefficient of -7.21 with a t statistic of -8.80 in a univariate regression specification, indicating that the stock's past performance negatively predicts its one-month-ahead price boost probability. Therefore, loser stocks have a much higher probability of having an extremely high daily return in the next month. Columns 2 to 9 augment the univariate regression by adding extra firm-specific attributes to the independent variables one at a time. The coefficients of PFM are estimated in the range of -6.99 and -7.66 in these specifications, and they are all significantly negative with t statistics between -8.92 and -9.14. Regression 9, which controls for all firm characteristics and risk attributes, shows that the slope coefficient of PFM is negative and highly significant, with a value of -7.21 and a t statistic of -9.14.

Table 5, Panel B reports the results of a cross-sectional regression where EXTREME denotes the indicator of price crash. The first column shows that PFM has a non-significantly positive coefficient of 0.64 with a t statistic of 1.66 in a univariate regression specification. Columns 2 to 9 augment the univariate regression by adding extra firm-specific attributes to the independent variables one at a time. The coefficients of PFM are in the range of 0.32 and 0.70 in these specifications, and they are all non-significantly positive with t statistics between 0.67 and 1.80. Therefore, the stock's past performance has no predicting power for the one-month-head price crash probability.

These results confirm my conjecture that loser stocks have a high probability of generating future extreme high returns given their optionality. In other words, loser stocks whose prices have suffered a severe decline may enjoy vigorous rebounds. Meanwhile, their prices have limited space to drop dramatically. Such attributes of loser stocks may attract investors with gambling preferences.

4.2.2. Portfolio analysis

To compare proportions of realized extremely high returns generated by loser and winner stocks further, I calculate the portfolio transition matrix. More specifically, I first sort all stocks in the sample into quintiles based on PFM calculated in month t-1 to produce losers (Quintile 1) and winners (Quintile 5). Then, I independently sort all stocks into quintiles by their MAX in month t.[9] The intersections of the five PFM groups and the five MAX groups generate 15 portfolios.

[9] Here, I apply MAX in the month after portfolio formation to represent the realized extremely high return. The results are quantitatively similar when I use the average of the five highest daily returns to conduct the analysis.

Table 6 reports the time-series means of the portfolio transition matrices. Specifically, the entry (i, j) in Table 6 represents that a stock in quintile i ranked by PFM (defined by the columns) will be in quintile j ranked by MAX (defined by the rows) in the next month. If the PFM in month t-1 and the MAX in month t are completely independent, then all the probabilities should be approximately 20%.[10] Instead, Table 6 shows that loser stocks in month t-1 are much more likely to have extremely high returns in month t. More specifically, Table 6 shows that stocks in Quintile 1 (loser) portfolio ranked on PFM in month t-1 have a 36.36% (t = 14.29) chance of appearing in the Quintile 5 (high-MAX) portfolio sorted by MAX in month t. Furthermore, the loser stocks have a very low probability of 9.91% (t = -14.25) of landing in the low-MAX portfolio in month t. In contrast, the winner stocks have roughly equal probabilities of appearing in each quintile portfolio sorted by MAX in month t, e.g., the winner stocks have 16.31% (22.28%) probability of landing in the low-MAX (high-MAX) portfolio in month t. This result further indicates that loser stocks seem more attractive to lotto investors, as they have a better chance of generating huge payoffs in the future than winner stocks.

4.2.3. Effective lottery-like features

Based on my results from cross-sectional and portfolio analyses, I next show the typical lottery-like features that depend greatly on stocks' past performances.

First, I construct 25 portfolios in month t-1 by independently sorting stocks based on MAX and PFM into quintile portfolios. Portfolio (1, 1) consists of stocks with low-MAX and low-PFM (loser), whereas portfolio (5, 5) is formed by stocks with high-MAX and high-PFM (winner). Next, like the analysis in Section 4.2.2, I apply MAX in month t to represent the realized extremely high return and I sort all stocks into quintile portfolios. Finally, I calculate the proportion of stocks in each of the 25 portfolios formed in month t-1 that transformed into high-MAX portfolios in month t.

Table 7 presents the time-series mean of probabilities to land in the high-MAX group in month t for each portfolio formed in month t-1. For example, entry (5, 1) in Table 7 shows that high-MAX stocks with poor performance have a probability of 61.53% of generating a maximum daily return

[10]The Newey and West (1987) t statistics statistics reported in parentheses measure the significance of any transition probability higher than 20%.

ranked in the top quintile in the next month, whereas entry (5, 5) shows that high-MAX stocks with good performance only have a 40.17% chance of landing in high-MAX portfolios. The last column reports the differences between the probabilities of loser and winner stocks, and their corresponding t statistics. Importantly, the probability difference between high-MAX stocks that performed poorly in the past three months and those that performed well is 21.36%, with a significant t statistic of 29.68. Thus, the lottery-like features depend greatly on the stock's past performance. In other words, the effectiveness of lottery-like measurements can be significantly improved by incorporating information on stocks' past performance. As a result, lotto investors prefer to gamble on lottery stocks with a poor performance, given their highest probability of winning.

4.3. Investors' Biased Beliefs

In addition to the effective lottery-like features of loser stocks with high-MAX, lotto investors may prefer to buy this type of stocks as they believe high-MAX stocks whose prices have suffered a severe decline have more space to jump upwards. In other words, lotto investors tend to believe that lottery-like stocks with poor performance may enjoy a vigorous rebound shortly, while those with good performance may not generate an extremely positive return given their current high prices. On the other hand, loser stocks without lottery-like features are likely to continue to perform poorly. As a result, an overly optimistic (pessimistic) belief about stocks with (without) lottery-like features results in a pronounced lottery premium among loser stocks.

4.3.1. Changes in Retail Holdings

Kumar (2009), Han and Kumar (2013), and Kumar et al. (2016) provide evidence that investors with a stronger propensity to gamble tend to be more attracted to lottery stocks. Bali et al. (2019) and Lin and Liu (2017) also find that the lottery premium is significantly greater among individual investors than institutional investors. Thus, I typically use retail investors to represent lotto investors. In this section, I examine the resulting theory that RHLD should be higher in loser stocks and especially in loser stocks with lottery-like features. Further, retail holdings of poorly performing lottery-like stocks should have had a significant increase in the previous quarter.

Panel A in Table 8 reports the result of univariate portfolio sort based on PFM. I sort all stocks in the sample into quintiles. For the stocks in each of these quintiles, I report the level (RHLD) and changes (ΔRHLD) in the number of retail holdings. I report the level (in month t-1) and changes over the months t-3 to t-1 based on the latest information during the quarter.

As shown in Panel A in Table 8, there appear to be more retail investors holding stocks that experienced a loss or gain, as I observe a higher number of RHLD in loser and winner stocks. More importantly, loser stocks have the largest RHLD. The last two rows report LMW. RHLD in loser stocks is significantly higher than that in winner stocks, with a difference of 6.20% ($t = 9.34$). The second column presents ΔRHLD for each quintile. I notice that RHLD is the largest in loser stocks, and, more importantly, that ΔRHLD is significantly positive (0.57%) in Quintile 1 (loser) and it monotonically decreases to -1.28% in Quintile 5 (winner). LMW is 1.85% with a t statistic of 11.31.

Next, I conduct conditional bivariate sorts as described in section 2.3. Panel B in Table8 reports the average RHLD for each portfolio. The last column confirms that RHLD is higher in loser stocks than in winner stocks, as all LMWs are significantly positive. The last two rows report the differences between high-MAX portfolios and low-MAX portfolios. As expected, RHLD in the high-MAX portfolio is significantly higher than in the low-MAX portfolio. More specifically, the values of high-low are in the range of 13.17% and 22.87%, and they are all significantly positive with t statistics between 9.40 and 12.88. I also notice that the highest value of RHLD appears in the bottom-left corner portfolio that contains stocks with lowest PFM and the highest MAX.

Panel C in Table 8 presents ΔRHLD for each portfolio. I observe that the ΔRHLDs are all negative except those in the loser column, indicating that the retail holdings experienced an increase only among poorly performing stocks. The last column shows a similar result to Panel A in Table 8 in that ΔRHLD in loser stocks is larger than in winner stocks. The last two rows show that ΔRHLD in the high-MAX portfolio is significantly positive only in the loser group. The most striking observation is that ΔRHLD has the largest value in the bottom-left corner portfolio at 1.01%. Therefore, during the previous quarter, loser stocks with the most lottery-like features experienced a significant increase in retail holdings, supporting the conjecture that investors who have gambling preferences are fascinated by loser stocks with lottery-like payoffs.

4.3.2. Lottery Premium, Earnings Announcements, and Analyst Forecasts

The biased beliefs framework proposed by Engelberg et al. (2018) suggests that investors may be overly optimistic (pessimistic) about certain groups of firms. The arrival of news may then force them to update their biased beliefs rapidly, resulting in abnormal returns over a short event window during which expected returns are close to zero. To test my conjecture that investors' biased beliefs produce the pronounced lottery premium in loser stocks further, I study the market reaction to firm-level earnings announcements. Abnormal announcement returns are defined as the difference between the actual buy-and-hold return over the event days $t = 1$ to $t = 1$ and the expected Fama and French (1993) three-factor buy-and-hold return.[11] I employ the methodology described in section 2.3. Specifically, in portfolio formation month t-1, I perform conditional bivariate portfolio sorts based on PFM and MAX and then I calculate value-weighted portfolio abnormal returns in next month.

Panel A in Table 9 reports the abnormal returns for each of the 25 portfolios and their corresponding Newey and West (1987) t statistics (in parentheses). The first (fifth) column presents the portfolios' abnormal returns in loser (winner) stocks. The low-MAX stocks in the loser group on average generate a significantly positive three-day abnormal return of 0.59% ($t = 3.60$), which is also significantly larger than the abnormal return of low-MAX stocks in winner stocks (0.17% with a t statistic of 1.42). Therefore, investors are overly pessimistic about loser stocks without lottery-like features, as they believe such type of stocks may continue to perform poorly. After receiving news about these stocks, investors are forced to update their pessimistic beliefs and thus to push up the stocks' prices. Importantly, the high-MAX stocks in the loser group on average have a significantly negative abnormal returns of -0.58% ($t = 2.75$), whereas the high-MAX stocks in the winner group yield abnormal returns of essentially zero with a t statistic of 0.01. Furthermore, the last column shows that in the high-MAX quintile, the return spread between loser stocks and winner stocks is -0.58% with a t statistic of -2.40. Therefore, investors who have gambling preference are essentially biased towards loser stocks, as they appear to overstate the prospects of lottery-like stocks that have performed poorly in the past.

The last two rows report the return spreads between high-MAX stocks and low-MAX stocks for each PFM quintile. I observe that only the return spread in the loser group is statistically significant with a spread of -1.16% and a t statistic of -3.92, whereas the return spread in the

[11] I obtain earnings announcement dates from the Compustat quarterly database. The data sample covers 1979 to 2017. Following Engelberg et al. (2018), I examine the firm's trading volume scaled by market trading volume for the day before, of, and after the reported earnings announcement date. I define the day with the highest volume as the earnings announcement day.

winner group is not significant (-0.17%, t = 0.73). In sum, these results suggest that investors' biased beliefs caused by gambling preference appear to apply only to loser stocks. Furthermore, investors are overly optimistic (pessimistic) about stocks with a high (low) maximum daily return, resulting in a highly significant negative (positive) abnormal three-day event-time return.

Panel B further presents Fama-MacBeth regressions of abnormal earnings announcements returns on the LOSER indicator that takes a value of one if the stock belongs to the bottom quintile based on its past performance, and zero otherwise, lottery-like features proxy (MAX or LTRY), and their interaction. I also control for a battery of variables, namely BETA, SIZE, MOM, REV, ILLIQ, IVOL, RSKEW, RKURT, CSK, and CKT. For both regressions, the coefficients of interaction term Proxy $\times$ LOSER are significantly negative. Importantly, in the regression where MAX is as the lottery-like feature, the coefficient of Proxy is not significant. Although the coefficient of Proxy (= LTRY) is significantly negative, its magnitude is ten times smaller than the coefficient of the interaction term. These findings support the analysis that lottery premium is induced by investors' biased beliefs, and it mainly affects stocks with poor performance.

I further measure investor expectation by analyst forecast. Bali, Brown, Murray, and Tang (2017) find that analyst coverage attracts retail investors' attention and it causes a strong lottery premium. Therefore, I suppose that analyst forecasts can affect and thus partially reflect retail investors' expectations. I investigate analyst forecast errors to examine whether investors' biased beliefs related to gambling preference are strong among poorly performing stocks. Analyst error for an announcement is the median forecast minus actual earnings, scaled by stock price in the previous month. A positive forecast error suggests over-optimism, and a negative error suggests investor over-pessimism. The right part of Panel B in 9 reports the results of regressing analyst forecast errors on LOSER indicator, lottery-like features (Proxy), Proxy $\times$ LOSER, and a battery of control variables. The coefficients of the interaction term Proxy $\times$ LOSER are all positively significant, while the coefficients of Proxy are not significant.

To summarize, I find evidence that lottery premium is induced by investors' biased belief, which is highly conditioned by stocks' past performance, and which mainly affects loser stocks.

4.3.3. Short-Sale Constraints and Optimistic Beliefs

The previous section pointed out that the significant lottery premium in loser stocks is driven by both the underperformance of high-MAX stocks and the outperformance of low-MAX stocks, and that high-MAX stocks account for most of the premium. I note that the short-sale constraint is the necessary condition for optimistic beliefs to induce overvaluation. It is then natural to expect an optimistic belief of gambling investors to induce overpricing only among stocks with binding short-sale constraints. For stocks without short-sale constraints, overpricing is unlikely, and therefore, we should not expect underperformance for high-MAX stocks. To test this prediction, I employ the same methodology as before, but with stocks that have options.[12]

As shown in Table 10, the lottery premium in the loser group experiences a significant decrease, with an alpha spread of -1.06% and a t statistic of -2.55, and the lottery premiums in the other PFM groups are all non-significant. Most importantly, the alpha of the high-MAX stocks in the loser group is no more significant with a value of -0.3% and a t statistic of -0.89. As expected, the alpha of the low-MAX stocks in the loser group remains significant, with a value of 0.76 and a t statistic of 3.20. Therefore, the significant lottery premium for loser stocks that have options is mainly driven by the outperformance of low-MAX stocks. In sum, the optimistic belief about high-MAX stocks with poor performance, combined with short-sale constraints, drives the short side of the lottery premium, whereas pessimistic beliefs about low-MAX loser stocks contribute to the long side of the lottery premium.

4.3.4. Investor Sentiment and Lottery Premium

Baker and Wurgler (2004) find that investor sentiment explains the cross-section of stock returns, with high sentiment being a significant predictor of the returns of more speculative stocks such as those of small firms, new firms, and highly volatile firms. When sentiment is high, subsequent returns on these stocks tend to be low and vice versa. Stambaugh et al. (2012) document that investor bias and mispricing is stronger during periods of high investor sentiment. Therefore, if investor biased belief applies to loser stocks, we should observe that the lottery premium is notably stronger in periods of high investor sentiment. In other words, as past performance is a typical stock

[12] Stocks that have options are less likely to face short-sale constraints. As I only have options data starting from 1996, I perform the portfolio sorts with the data sample from 1996 to 2017.

characteristic that draws the attention of lotto investors, the difference in lottery premium between loser and winner stocks should be much higher when investors have more propensity to gamble (high sentiment periods). Furthermore, lotto investors are more likely to overvalue loser stocks with lottery-like features, while disliking those without lottery-like features. More specifically, I conjecture that when sentiment is high, investors are more optimistic (pessimistic) about the future payoffs of loser stocks with (without) lottery-like features than when sentiment is low.

I examine the relationship between sentiment and the lottery premium using the Baker and Wurgler (2004) (BW) sentiment index which is the first principal component of six underlying sentiment proxies: closed-end fund discount, the number and the first-day returns of IPOs, NYSE turnover, the equity share in total new issues, and the dividend premium. Following Baker and Wurgler (2004) and Stambaugh et al. (2012), I define a high (low) sentiment month as one in which the BW index is above (below) the sample median value. In each sentiment period, I perform conditional bivariate sorts as described in Section 2.3. Specifically, I first perform a sequential sort by creating portfolios ranked by PFM using breakpoints of 30%-40%-30%. Then, within each PFM portfolio, I form a second set of quintile portfolios ranked on stock lottery-like features (MAX or LTRY).

Panel A in Table 11 shows that during periods of high sentiment, the lottery premium for loser stocks is highly significant with an alpha spread of -3.04% and a t statistic of -7.98. However, during periods of low sentiment, the lottery premium for loser stocks is only -0.79% with a t statistic of -2.16. Although the lottery premium for loser stocks is still marginally significant, it experiences a significant decline when I move from high sentiment periods to low sentiment periods. Furthermore, the lottery premiums for winner stocks are quite similar during high and low sentiment periods (-0.53% and -0.52%), and they are both non-significant. What is more, the difference in lottery premium between loser and winner stocks is relatively small when investor sentiment is low. In addition, I find that the significant difference in alpha spread during different sentiment periods is mainly driven by high-MAX stocks, as the alpha of high-MAX stocks in the loser group during high sentiment periods is much lower than that during low sentiment periods (-2.25% vs. -0.78%). This is consistent with Miller (1977) argument that due to short-sale constraints, the price of a stock reflects the views of the most optimistic investors. Importantly, I also observe that the alpha of low-MAX stocks in the loser group during high sentiment periods is much higher than that during low sentiment periods (0.79% vs. 0.22%), indicating that loser stocks with low-MAX are more likely to be underpriced during high sentiment periods. These results show that lotto investors' biased beliefs mainly apply to loser stocks.

As shown in Panel B in Table 11, similar results occur when I use LTRY as the lottery-like feature proxy. Specifically, the lottery premium for loser stocks is highly significant with an alpha spread of -2.02% and a t statistic of -7.52, whereas, during periods of low sentiment, the lottery premium for loser stocks is only -0.55% with a t statistic of -1.96. Furthermore, the alpha of high-MAX (low-MAX) stocks in the loser group is much lower (higher) during high sentiment periods. In addition, all alphas for winner portfolios are non-significant in both sentiment periods. Overall, when sentiment is high, investors are more optimistic (pessimistic) about the future payoffs of loser stocks with (without) lottery-like features than when sentiment is low.

5. Cross-Sectional Regression Analysis

The portfolio sorts present strong evidence that the lottery premium is significantly pronounced in loser stocks. Loser stocks with lottery-like features significantly underperform loser stocks without lottery-like payoffs. While the portfolio analysis is nonparametric in that I do not impose a functional form on the relation between PFM, lottery-like features, and future returns, it does not allow us to account for other control variables jointly. The extant literature highlights the role of specific firm characteristics in predicting future returns. Therefore, it is essential to investigate how PFM affects lottery premium after controlling for the effects of other firm characteristics.

This section complements my portfolio-level analysis by performing Fama and MacBeth (1973) cross-sectional regressions using individual stock returns. Specifically, I estimate the following cross-sectional regression in each month:

$$\begin{aligned} R_{i,t} = &\lambda_{i,0} + \lambda_{i,1}\text{LOSER}_{i,t-1} + \lambda_{i,2}\text{PROXY}_{i,t-1} \\ &+ \lambda_{i,3}\text{PROXY}_{i,t-1} \times \text{LOSER}_{i,t-1} + \lambda_{i,4}X_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{7}$$

where $R_{i,t}$ is the realized excess return on stock i in month t, $\text{LOSER}_{i,t-1}$ is a dummy variable that takes a value of one in month t-1 if the stock belongs to the bottom quintile based on PFM, and zero otherwise, $\text{PROXY}_{i,t-1}$ denotes the proxy for stock i's lottery-like payoffs (MAX or LTRY) in month t-1, and $X_{i,t-1}$ is a vector of control variables for stock i in month t-1, including BETA, SIZE, CGO, MOM, REV, ILLIQ, IVOL, RSKEW, RKURT, CSK, and CKT.[13] The regression specification also allows for interaction of the stock's past performance indicator $\text{LOSER}_{i,t-1}$ and lottery-like features proxy $\text{PROXY}_{i,t-1}$ to investigate their dependent effects on future returns

I report the time-series averages (and the associated t statistics) of the monthly cross-sectional regression coefficients in Table 12. Panel A in Table 12 presents the regression results estimated using the ordinary least squares (OLS) method and Panel B reports the results estimated using the weighted least squares (WLS) methodology following Asparouhova, Bessembinder, and Kalcheva (2013) where each observed return is weighted by one plus the observed prior return on the stock.

The first two columns in Panel A report the estimation results from the regressions where MAX is the lottery-like feature proxy. The first column reports the coefficient estimates of my interest variables (LOSER, PROXY, and their interactions) for a baseline model that does not include other control variables. This model is used to examine my main finding that lottery premium is highly dependent on the stock's past performance. The results show that the LOSER coefficient estimate is significantly negative, indicating a momentum effect that loser stocks continue to perform poorly in general. Importantly, the interaction term (PROXY $\times$ LOSER) is significantly negative and it subsumes the effect of PROXY, supporting the result of portfolio analysis that the lottery premium is significantly pronounced in loser stocks. More specifically, the average slope, λ_3, from the monthly regressions is -18.62 with a t statistic of -7.13, implying that a one standard deviation increase in MAX in the previous month translates to a current month annualized return that is 4.25% lower for loser stocks.[14] Furthermore, the return spread in MAX between Quintiles 5 and 1 is 8.16% ($= 9.95\% - 1.79\%$) in the loser group. Multiplying this spread by the average slope yields an estimate of the monthly lottery premium of 1.52%. As shown in the second column, after incorporating all control variables simultaneously, the slope coefficient of the interaction term increases slightly and remains negative and statistically significant, 13.01% (t = 5.62). These results show that MAX combined with the loser indicator has distinct, significant information orthogonal to BETA, SIZE, past returns, LIILQ, IVOL, CSK, CKT, RSKEW, RKURT and, importantly, CGO. Similar results

[13] Given that IVOL is a component of LTRY, I exclude IVOL from this specification when lottery payoffs are proxied by LTRY to avoid multicollinearity.

[14] The time-series average of standard deviation of MAX is 0.019. Therefore, the annualized change in stock return corresponding to a one standard deviation increase in MAX is 4.25% ($0.019 \times -18.62\% \times 12$).

are observed in Panel B for the WLS regressions.

As expected, I get similar results when I use LTRY as the lottery-like feature proxy. Specifically, the average coefficients of the interaction term in the two model specifications are -0.66 and -0.36, respectively. They are both statistically significant as reported in the third and fourth columns. In addition, the monthly lottery premium in loser stocks calculated using the slop coefficient, λ_3, in the third column is 0.79%, which is only half of the lottery premium obtained with MAX as the lottery-like feature proxy, but it is still economically significant. These results also apply to the WLS estimates.

In summary, I find that the lottery premium is strongly affected by stocks' past performance by using firm-level cross-sectional regressions, further confirming the results obtained from the portfolio sorts analysis. The evidence confirms that the stock's past performance is a significant factor in the return spread caused by lotto investor biased beliefs, even after controlling for other determinants of asset returns. Therefore, my results have theoretical and practical implications for asset pricing and portfolio management. Existing theories of stock gambling such as the optimal expectations model of Brunnermeier, Gollier, and Parker (2007) argue that investors are drawn to lottery-like securities because they are optimistic about the future payoffs of these securities. I provide further evidence that the lottery premium is strongly dependent on stocks' past performance.

6. Robustness Tests

6.1. Subperiod Analysis

In Table 13, I repeat the analysis for two different subperiods (1965-1996 and 1996-2017). The results show that all inferences remain stable.[15] As expected, the lottery premium in the loser group is higher during the first period, with an alpha spread of -2.65% (t = 10.34). Notably, the alpha spread between high-MAX stocks and low-MAX stocks in the loser group remains significantly negative during the most recent period (-1.92%, t = 4.22). Importantly, the lottery premiums in the winner group are much lower than those in the loser group during both periods, emphasizing the significant role of stock past performance in determining the lottery premium.

[15] I only report the result of the portfolio sort based on MAX, as I have a similar result for the portfolio sort ranked by LTRY.

6.2. *Price Restrictions and Lottery premium*

In this section, I explore whether the lottery premium pattern is robust to different sample specifications. Jiang, Xu, and Yao (2009) document that a $5 price restriction helps in avoiding market microstructure-related issues.

Panel A of Table 14 presents the FFCPS alpha for each of the 15 portfolios created by conditional double sorts when a $5 price restriction is imposed. I continue to find that the lottery premium is very high among stocks with low past returns (losers) but very low among stocks with high past returns (winners). Specifically, the alpha spread is -1.62% per month with a t statistic of -5.34 for loser stocks, whereas the alpha spreads are -0.54% (t = 2.03) and -0.64% (t = 2.26) for medium and winner stocks, respectively. Furthermore, the lottery premium for loser stocks is driven by both the outperformance of low-MAX stocks and the underperformance of high-MAX stocks, as the loser stocks with low-MAX have a significant positive alpha of 0.59% per month (t = 5.07), and those with high-MAX have a significant negative alpha of -1.03% per month (t = 4.53). Besides, low-MAX stocks with poor performance have a higher alpha than those with good performance as the LMW alpha spread in the low-MAX group is 0.39% (t = 2.46). Furthermore, the high-MAX stocks that perform poorly have much lower alphas than those that perform well, since the LMW alpha spread in the high-MAX group is -0.59% (t = 2.77). Consequently, the LMW alpha spread (-0.98%) in the high-low portfolio is both economically and statistically significant (t = 3.13).

Similar results are obtained when lottery-like payoffs are proxied by LTRY. Panel B of Table 14 shows that in the loser group, the FFCPS alpha spread is -1.22% per month with a t-statistic of -4.52. In contrast, in the winner group, the FFCPS spread is only -0.10% with an insignificant t-statistic of -0.64. The LMW alpha spread (-1.12%) in High-Low portfolio remains both economically and statistically significant ($t = -3.83$). In addition, the lose stocks with Low-MAX have a significant positive alpha of 0.42% per month (t = 4.13), and those with High-MAX have a significant negative alpha of -0.80% per month ($t = -3.16$).

Overall, these results are consistent with the previous findings that the lottery premium is significantly pronounced among loser stocks and that the lottery premium is driven by both the outperformance of loser stocks without lottery-like features and the underperformance of those with lottery-like features. Furthermore, price restriction decreases the lottery premium by excluding low-priced stocks, which are more likely to have lottery-like features and thus be overpriced.

6.3. Alternative Measures of Stocks' Past Performance

In this section, I explore whether the lottery premium pattern is robust to alternative measurements of PFM. In my primary analysis, I employ a three-month cumulative return as the proxy for stock past performance. Instead, I can use long-term performance measures such as six- or twelve-month cumulative return.

Panel A of Table 15 presents the FFCPS alpha for each of the 15 portfolios created by conditional double sorts when I use twelve-month cumulative return and MAX as the proxies for stock's past performance and lottery-like feature, respectively.[16] e continue to find that the lottery premium is very high among stocks with low past returns (losers), but very low among stocks with high past returns (winners). Specifically, the alpha spread is -2.23% per month with a t statistic of -8.63 for loser stocks, whereas the alpha spreads are -0.51% (-3.59 t statistic) and -0.54% (-2.49 t statistic) for medium and winner stocks, respectively. The lottery premium is only slightly larger than one calculated by using a three-month cumulative return as the proxy for PFM (-2.13%), and all the other earlier patterns remain.

Similar results are obtained when a six-month cumulative return denotes PFM. Panel B of Table 15 shows that in the loser group, the FFCPS alpha spread is -2.23% per month with a t statistic of -8.55. In contrast, in the winner group, the FFCPS spread is only -0.57% with a non-significant t statistic of -2.50. Therefore, my findings are robust to alternative measurements of PFM, and the results obtained by using long-term cumulative returns are slightly more robust.

6.4. Persistence of the Lottery Premium

I investigate the long-term lottery premium over the following six months by constructing portfolios with overlapping holding periods. Table 16 reports the FFCPS monthly alphas for portfolios in loser and winner groups for one- to six-month holding periods.[17] The risk-adjusted return differences between the high-MAX and low-MAX portfolios in the loser group are statistically significant up to four-month holding periods. Specifically, the t+2 through t+4 average monthly

[16] In an unreported table, I find similar results using LTRY as the proxy for lottery-like features.

[17] I only report results based on portfolios formed using MAX as the lottery-like features. The results are quantitatively similar when I use LTRY to represent lottery-like features.

alpha spreads are -1.43% ($t = 7.00$), -0.97% ($t = 4.98$) and -0.59% ($t = 2.93$), respectively. There is no evidence of a lottery premium in months t+2 through t+6 for winner stocks. Thus, the lottery premium for loser stocks is persistent and concentrated within the following four months.

7. Conclusion

In this paper, I shed new light on the cross-sectional variation in lottery premium induced by investors' gambling preference. Specifically, I find that the lottery premium is concentrated on stocks that performed poorly in the previous months from t-3 to t-1. I show that the finding is related to the option-like feature of loser stocks that attract the investors who have gambling preferences, typically retail investors. I find that PFM negatively predicts a one-month-ahead price boost probability. Therefore, the prices of loser stocks have a higher probability of jumping upwards than those of winner stocks. More detailed analysis reveals that lottery-like stocks with poor performance have a much higher chance of generating a huge payoff in the next month than those with excellent performance. Thus, the lottery-like features depend strongly on the stock's past performance. In other words, the effectiveness of lottery-like measurements can be significantly improved by incorporating information on the stock's past performance. As a result, lotto investors prefer to gamble on lottery stocks with a poor performance given their highest probability of winning and relatively low costs.

Lotto investors may prefer to buy such types of stocks because they believe stocks whose prices have suffered a severe decline have more space for their prices to jump upwards. In other words, lotto investors tend to believe that lottery-like stocks with poor performance may enjoy a vigorous rebound shortly, while those with excellent performance may have a lower probability of generating an extremely positive return given their current high prices. On the other hand, loser stocks without lottery-like features may continue to perform poorly. As a result, overly optimistic (pessimistic) beliefs about stocks with (without) lottery-like features results in a pronounced lottery premium among loser stocks.

I find that retail holdings are much higher in loser stocks than in winner stocks. Especially, in the previous quarter, loser stocks with lottery-like features experienced a significant increase in retail holdings, suggesting that investors who have gambling preference are fascinated by loser stocks with most lottery-like payoffs. Furthermore, I discover that investors' biased beliefs appear to apply only to loser stocks. The arrival of news in the following month forces investors to update their optimistic (pessimistic) beliefs about loser stocks with (without) lottery-like features rapidly, and it results in highly significant negative (positive) abnormal three-day event-time returns. Thus, the lottery premium for loser stocks is driven by both the underperformance of loser stocks with lottery-like features and the outperformance of those without lottery-like features. Consistent with prior findings, I show that the optimistic belief about loser stocks with lottery-like features, combined with short-sale constraint, drives the short side of the lottery premium. Finally, I demonstrate that the lottery premium is stronger following high investor sentiment states, and, importantly, that the difference in lottery premium between loser and winner stocks is small when sentiment is low. Thus, the loser effect only exists when investors have a high propensity to gamble.

My analysis suggests several avenues for future research. First, it is not clear whether other stock's characteristics can also be applied to improve the effectiveness of the measurement of lottery-like features. Second, it is not clear how the results vary across different assets and countries. Third, it may also be of interest to ascertain whether the results apply in derivatives markets, as lotto investors may prefer to gamble on derivatives given their possible high payoffs.

References

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

An, Li, Huijun Wang, Jian Wang, and Jianfeng Yu, forthcoming, Lottery-related anomalies: the role of reference-dependent preferences, *Management Science* .

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *The Journal of Finance* 61, 259–299.

Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva, 2013, Noisy prices and inference regarding returns, *The Journal of Finance* 68, 665–714.

Baker, Malcolm, and Jeffrey Wurgler, 2004, Investor sentiment and the cross-section of stock returns, *The Journal of Finance* 61, 1645–1680.

Bali, Turan G., Stephen J. Brown, Scott Murray, and Yi Tang, 2017, A lottery-demand-based explanation of the beta anomaly, *Journal of Financial and Quantitative Analysis* 52, 2369–2397.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427–446.

Bali, Turan G., David Hirshleifer, Lin Peng, and Yi Tang, 2019, Attention, social interaction, and investor attraction to lottery stocks, Working paper.

Baltussen, Guido, Sjoerd van Bekkum, and Bart van der Grient, 2018, Unknown unknowns: uncertainty about risk and stock returns, *The Journal of Financial and Quantitative Analysis* 53, 1615–1651.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: the implications of probability weighting for security prices, *The American Economic Review* 98, 2066–2100.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect theory and stock returns: an empirical test, *The Review of Financial Studies* 29, 3068–3107.

Bernard, Victor L., and Jacob K. Thomas, 1989, Post-earnings-announcement drift: delayed price response or risk premium?, *Journal of Accounting Research* 27, 1–36.

Boyer, Brian H., and Keith Vorkink, 2014, Stock options as lotteries, *The Journal of Finance* 4, 1485–1527.

Brunnermeier, Markus K., Christian Gollier, and Jonathan A. Parker, 2007, Optimal beliefs, asset prices, and the preference for skewed returns, *The American Economic Review* 97, 159–165.

Brunnermeier, Markus K., and Jonathan A. Parker, 2005, Optimal expectations, *The American Economic Review* 95, 1092–111.

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

Cremers, K. J. Martijn, and Vinay B. Nair, 2005, Governance mechanisms and equity prices, *The Journal of Finance* 60, 2859–2894.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *The Journal of Finance* 66, 1461–1499.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221–247.

DellaVigna, Stefano, and Joshua M. Pollet, 2009, Investor inattention and friday earnings announcements, *The Journal of Finance* 64, 709–749.

Engelberg, Joseph, R. David Mclean, and Jeffrey Pontiff, 2018, Anomalies and news, *The Journal of Finance* 73, 1971–2001.

Eraker, Bjrn, and Mark Ready, 2015, Do investors overpay for stocks with lottery-like payoffs? an examination of the returns of otc stocks, *Journal of Financial Economics* 115, 486–504.

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

Fama, Eugene F, and Kenneth R French, 2008, Dissecting anomalies, *The Journal of Finance* 63, 1653–1678.

Fama, Eugene F, and James D MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *The Journal of Political Economy* 81, 607–636.

Han, Bing, and Alok Kumar, 2013, Speculative retail trading and asset prices, *Journal of Financial and Quantitative Analysis* 48, 377–404.

Harvey, Campbell R, and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *The Journal of Finance* 55, 1263–1295.

Hutton, Amy P., Alan J. Marcus, and Hassan Tehranian, 2009, Opaque financial reports, r2, and crash risk, *Journal of Financial Economics* 94, 67–86.

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

Jiang, George J., Danielle Xu, and Tong Yao, 2009, The information content of idiosyncratic volatility, *Journal of Financial and Quantitative Analysise* 44, 128.

Kim, Jeong-Bon, Yinghua Li, and Liandong Zhang, 2011, Corporate tax avoidance and stock price crash risk: Firm-level analysis, *Journal of Financial Economics* 100, 639–662.

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

Kumar, Alok, Jeremy K. Page, and Oliver G. Spalt, 2016, Gambling and comovement, *Journal of Financial and Quantitative Analysis* 51, 85111.

Lin, Tse-Chun, and Xin Liu, 2017, Skewness, individual investor preference, and the cross-section of stock returns, *Review of Finance* 22, 1841–1876.

Lin Peng, Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.

Lo, Andrew W., and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, *The Review of Financial Studies* 3, 175–205.

Merton, Robert C., 1974, On the pricing of corporate debt: the risk structure of interest rates, *The Journal of Finance* 29, 449–470.

Miller, Edward M., 1977, Risk, uncertainty, and divergence of opinion, *The Journal of Finance* 32, 1151–1168.

Mitton, Todd, and Keith Vorkink, 2007, Equilibrium underdiversification and the preference for skewness, *The Review of Financial Studies* 20, 1255–1288.

Newey, Whitney K., and Kenneth D. West, 1987, Hypothesis testing with efficient method of moments estimation, *International Economic Review* 28, 777–787.

Peng, Lin, 2005, Learning with information capacity constraints, *Journal of Financial and Quantitative Analysis* 40, 307329.

Pstor, ubo, and RobertF Stambaugh, 2003, Liquidity risk and expected stock returns, *The Journal of Political Economy* 111, 642–685.

Routledge, Bryan R, and Stanley E Zin, 2010, Generalized disappointment aversion and asset prices, *The Journal of Finance* 65, 1303–1332.

Shumway, Tyler, 1997, The delisting bias in crsp data, *The Journal of Finance* 52, 327–340.

Shumway, Tyler, and Vincent A Warther, 1999, The delisting bias in crsp's nasdaq data and its implications for the size effect, *The Journal of Finance* 54, 2361–2379.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *The Journal of Finance* 70, 1903–1948.

Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.

Walkshusl, Christian, 2014, The max effect: European evidence, *Journal of Banking and Finance* 42, 1–10.

Zhang, X Frank, 2006, Information uncertainty and stock returns, *The Journal of Finance* 61, 105–137.

Zhong, Angel, and Philip Gray, 2016, The max effect: An exploration of risk and mispricing explanations, *Journal of Banking and Finance* 65, 76–90.

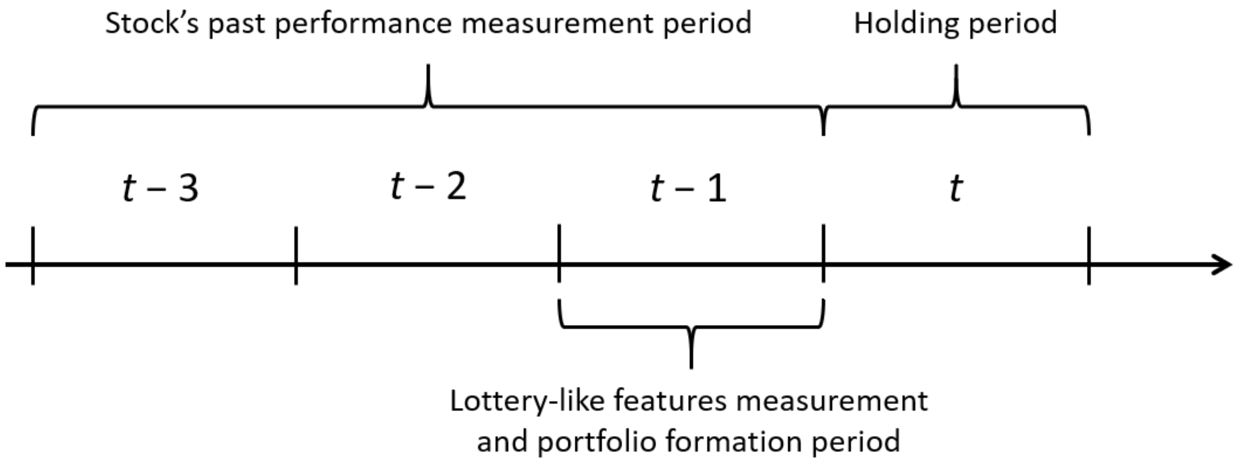

Fig. 1. This figure presents the time line of the three measurement periods I consider in this paper. The first period consists of 3 months, months t-3 to t-1, and it is used to measure stock's past performance (PFM). The second period (month t-1) is used as the formation month to sort stocks based on PFM and lottery-like features measured in the same month. The last period (month t) is used to calculate the average holding period returns for each portfolio.

Table 1: Univariate portfolio sort.
This table presents the value-weighted portfolio results for the period July 1965 to December 2017. Specifically, every month, I sort stocks based on their lottery-like features over the past month to form quintile portfolios. Quantile 1 is the portfolio with the lowest lottery-like features in the past month, and quintile 5 is the portfolio with the highest lottery-like features in the past month. Following Bali et al. (2019), I construct two types of proxies for lottery-like features. The first one is the maximum daily return of a stock in the previous month, denoted MAX (Bali et al. (2011). The second proxy for lottery-like payoffs denoted as LTRY is adapted from Kumar (2009). The first two columns reports the results of portfolios sorted by MAX, and the next two columns reports the results of portfolios sorted by LTRY. The first and third columns report the value-weighted average monthly excess returns for each quintile portfolio and the return difference between quintile 5 (High) and quintile 1 (Low). The second and fourth columns present the corresponding risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD) and liquidity factor (PS). Newey and West (1987) t statistics are given in parentheses.

	MAX		LTRY	
	Excess Return	FFCPS Alpha	Excess Return	FFCPS Alpha
Low	0.60	0.13	0.54	0.05
		(1.74)		(1.11)
2	0.59	0.05	0.55	0.02
		(1.27)		(0.59)
3	0.61	0.04	0.63	0.01
		(0.60)		(0.23)
4	0.49	-0.16	0.54	-0.06
		(-1.80)		(-0.73)
High	0.12	-0.59	0.20	-0.46
		(-2.78)		(-3.61)
High-Low	-0.47	-0.72	-0.34	-0.51
	(-1.71)	(-3.15)	(-1.62)	(-3.69)

Table 2: Stock's past performance and lottery premium.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts on stock's past performance defined as it's cumulative returns across months *t*-3 to *t*-1, and lottery-like features. Following Bali et al. (2019), I construct two types of proxies for lottery-like features. The first one is the maximum daily return of a stock in the previous month, denoted MAX (Bali et al. (2011). The second proxy for lottery-like payoffs denoted as LTRY is adapted from Kumar (2009). For each month, I group all stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending order of PFM. The bottom 30% stocks are denoted as losers, and the top 30% stocks are marked as winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending order of the lottery proxy (MAX or LTRY) calculated in month *t*-1. The intersections of the three PFM groups and the five lottery proxy groups generate total 15 value-weighted portfolios. Finally, I calculate portfolios' mean holding period returns in month *t*. Panel A reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD) and liquidity factor (PS) for each of the 15 portfolios when I use MAX as the lottery proxy. Panel B presents the FFCPS alphas of each portfolios with LTRY as the lottery proxy. Newey and West (1987) *t*-statistics are given in parentheses. LMW denotes loser minus winner.

Panel A: MAX as the lottery proxy

	Loser	Medium	Winner	LMW
Low	0.57	0.24	0.15	0.42
	(5.22)	(3.25)	(1.75)	(2.64)
2	0.18	0.10	-0.15	0.33
	(1.50)	(1.37)	(-1.57)	(1.86)
3	-0.23	0.09	-0.17	-0.05
	(-1.45)	(1.14)	(-1.47)	(-0.27)
4	-0.75	-0.18	-0.43	-0.32
	(-4.23)	(-1.69)	(-2.75)	(-1.27)
High	-1.56	-0.42	-0.52	-1.04
	(-6.64)	(-3.25)	(-2.33)	(-3.88)
High-Low	-2.13	-0.66	-0.67	-1.46
	(-7.73)	(-3.99)	(-2.70)	(-5.11)

Panel B: LTRY as the lottery proxy

	Loser	Medium	Winner	LMW
Low	0.29	0.11	-0.08	0.36
	(2.87)	(2.15)	(-0.94)	(2.25)
2	-0.17	0.03	-0.19	0.02
	(-1.40)	(0.44)	(-1.81)	(0.13)
3	-0.13	0.16	-0.25	0.11
	(-0.91)	(2.30)	(-1.87)	(0.56)
4	-0.58	-0.10	-0.31	-0.27
	(-3.96)	(-1.17)	(-2.27)	(-1.32)
High	-1.19	-0.15	-0.35	-0.84
	(-6.54)	(-1.16)	(-2.22)	(-4.29)
High-Low	-1.47	-0.26	-0.27	-1.20
	(-7.19)	(-1.88)	(-1.54)	(-5.31)

Table 3: Characteristics of portfolios sorted by stock's past performance (PFM).
This table reports the mean value of characteristics for each portfolio created by sorting on PFM. The characteristics include market beta (BETA), firm size (SIZE), book-to-market ratio (BM), return reversal (REV), illiquidity (ILLIQ), maximum daily return in previous month (MAX), idiosyncratic volatility (IVOL) and idiosyncratic skewness (ISKEW), price per share (PRC), and capital gain overhang (CGO).

	PFM	BETA	SIZE	BM	REV	ILLIQ	MAX	IVOL	ISKEW	PRC	CGO
Loser	-16.23	0.85	10.96	0.71	-0.16	-1.23	4.12	2.75	0.23	9.62	-0.30
Medium	1.03	0.73	11.88	0.76	0.24	-2.39	3.32	1.72	0.16	18.58	-0.05
Winner	21.08	0.82	11.70	0.73	-0.07	-2.45	3.56	2.07	0.18	17.33	0.07

Table 4: Trivariate portfolio sorts.

This table reports the results of trivariate portfolio sorts applied to control for various firm's characteristics. Specifically, I first perform a sequential sort by creating three portfolios with $30\% - 40\% - 30\%$ breakpoints ranked by control variables. Then, within each control variable portfolio, I group all stocks into three portfolios with $30\% - 40\% - 30\%$ breakpoints ranked by stock's past performance (PFM). Next, I group PFM portfolios that have the rank cross the control portfolios into three aggregate PFM portfolios. As a result, the aggregate PFM portfolios have a similar level of control variables. Finally, within each aggregate portfolios, I form a set of quintile portfolios ranked on stock's lottery-like features. The control variables include firm size (SIZE), price per share (PRC), return reversal (REV), illiquidity (ILLIQ), idiosyncratic volatility (IVOL) and capital gain overhang (CGO). Panel A (B) reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD) and liquidity factor (PS). Newey and West (1987) t statistics are given in parentheses. Average represents the mean level of control variables in each PFM portfolios.

Panel A: MAX as the lottery proxy

	SIZE		PRC		REV		ILLIQ		IVOL		CGO	
	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner
Average	11.68	11.74	13.92	14.82	0.00	0.00	-2.08	-2.20	2.07	2.03	-0.14	-0.12
Low	0.48	0.07	0.40	0.10	0.48	0.11	0.47	0.11	0.44	0.13	0.32	0.10
	(4.61)	(0.74)	(4.43)	(0.99)	(4.49)	(1.26)	(4.56)	(1.13)	(4.64)	(1.53)	(3.27)	(1.01)
2	0.35	-0.18	0.26	-0.17	0.23	-0.12	0.35	-0.15	0.29	-0.08	0.24	-0.16
	(3.27)	(-2.10)	(3.12)	(-1.91)	(2.07)	(-1.14)	(3.14)	(-1.58)	(3.03)	(-0.86)	(2.16)	(-1.74)
3	-0.19	-0.20	-0.14	-0.16	-0.24	-0.17	-0.17	-0.17	-0.07	-0.12	-0.05	-0.22
	(-1.39)	(-1.87)	(-1.18)	(-1.40)	(-1.65)	(-1.62)	(-1.23)	(-1.55)	(-0.58)	(-1.04)	(-0.38)	(-1.81)
4	-0.51	-0.30	-0.39	-0.33	-0.51	-0.31	-0.61	-0.29	-0.45	-0.21	-0.40	-0.35
	(-3.11)	(-1.94)	(-2.33)	(-2.07)	(-2.58)	(-2.07)	(-3.55)	(-1.83)	(-2.31)	(-1.64)	(-2.42)	(-1.89)
High	-1.37	-0.48	-1.38	-0.58	-1.56	-0.52	-1.34	-0.48	-1.38	-0.57	-1.21	-0.43
	(-6.41)	(-2.22)	(-6.43)	(-2.68)	(-6.78)	(-2.36)	(-6.35)	(-2.26)	(-6.12)	(-2.64)	(-4.47)	(-2.26)
High-Low	-1.85	-0.55	-1.78	-0.68	-2.04	-0.63	-1.81	-0.59	-1.82	-0.70	-1.53	-0.53
	(-7.25)	(-2.26)	(-7.09)	(-2.67)	(-7.24)	(-2.51)	(-7.06)	(-2.41)	(-7.12)	(-2.85)	(-5.52)	(-2.24)

Table 4: Continued.

Panel B: LTRY as the lottery proxy

	SIZE		PRC		REV		ILLIQ		IVOL		CGO	
	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner
Average	11.68	11.74	13.92	14.82	0.00	0.00	-2.08	-2.20	2.07	2.03	-0.14	-0.12
Low	0.37	-0.11	0.28	-0.09	0.33	-0.12	0.34	-0.08	0.34	-0.04	0.21	-0.06
	(4.17)	(-1.42)	(3.45)	(-1.01)	(3.52)	(-1.45)	(3.89)	(-0.95)	(3.83)	(-0.50)	(2.57)	(-0.77)
2	0.00	-0.27	0.06	-0.27	-0.06	-0.15	-0.05	-0.27	0.02	-0.18	-0.01	-0.27
	(0.03)	(-2.61)	(0.62)	(-2.66)	(-0.47)	(-1.51)	(-0.43)	(-2.46)	(0.24)	(-2.16)	(-0.13)	(-2.72)
3	-0.06	-0.17	-0.08	-0.20	-0.08	-0.17	0.01	-0.14	-0.17	-0.14	-0.08	-0.28
	(-0.55)	(-1.34)	(-0.60)	(-1.53)	(-0.62)	(-1.40)	(0.08)	(-1.09)	(-1.31)	(-1.25)	(-0.58)	(-2.10)
4	-0.30	-0.37	-0.33	-0.24	-0.45	-0.27	-0.30	-0.37	-0.30	-0.21	-0.13	-0.29
	(-1.90)	(-2.90)	(-2.24)	(-2.09)	(-3.20)	(-1.98)	(-1.88)	(-2.96)	(-1.94)	(-1.66)	(-0.82)	(-2.36)
High	-1.07	-0.15	-1.06	-0.20	-1.20	-0.22	-1.09	-0.19	-1.04	-0.25	-0.84	-0.16
	(-6.06)	(-0.88)	(-6.35)	(-1.20)	(-6.17)	(-1.41)	(-6.12)	(-1.18)	(-6.05)	(-1.48)	(-4.84)	(-0.90)
High-Low	-1.44	-0.03	-1.33	-0.11	-1.52	-0.11	-1.43	-0.11	-1.39	-0.21	-1.05	-0.10
	(-7.31)	(-0.18)	(-7.40)	(-0.62)	(-7.18)	(-0.60)	(-7.14)	(-0.61)	(-7.37)	(-1.19)	(-5.68)	(-0.49)

Table 5: Prediction of future extreme events.
This table reports the estimated results from cross-sectional regressions to examine the predictive power of firm's characteristics to future extreme events. I first define the monthly price boost (crash) measure as the indicator variable that equals one for a firm-month that experience one or more price boost (crash) days during the month, and zero otherwise. Similar to the definition used in Hutton et al. (2009) and Kim et al. (2011), price boost (crash) days in a given month are days in which the firm experiences daily returns are 3.09 (0.1% for normal distribution) standard deviations above (below) the mean daily returns over the entire year. I then perform the cross-sectional regression by first conducting a monthly logistic regression of the one-month-ahead price boost or crash measure on current PFM (cumulative returns across from months t-3 to t-1), controlling for a set of firm characteristics. I then aggregate coefficients across all months and compute the Newey and West (1987) t statistics (reported in parentheses). Panel A (B) reports the predictability of firm's characteristics to stock price boost (crash) probability.

Panel A: Dep. Var: Stock price boost probability

	1	2	3	4	5	6	7	8	9
Intercept	0.14	0.16	0.25	0.25	0.24	0.20	0.20	0.24	0.24
	(25.15)	(27.44)	(20.59)	(23.37)	(23.93)	(12.23)	(13.11)	(15.93)	(16.16)
PFM	-7.21	-7.66	-7.09	-7.05	-7.09	-7.01	-6.99	-7.18	-7.21
	(-8.80)	(-8.92)	(-9.04)	(-9.07)	(-9.06)	(-9.11)	(-9.06)	(-9.04)	(-9.14)
BETA		-1.89	-0.74	-0.76	-1.13	-0.91	-1.05	-1.44	-2.09
		(-6.23)	(-3.02)	(-3.20)	(-4.78)	(-4.16)	(-4.26)	(-6.10)	(-10.38)
SIZE			-0.79	-0.78	-0.68	-0.35	-0.36	-0.62	-0.71
			(-9.09)	(-9.59)	(-8.54)	(-2.03)	(-2.21)	(-4.01)	(-4.69)
BM				0.16	0.24	0.17	0.12	0.03	0.02
				(0.72)	(1.04)	(0.80)	(0.61)	(0.14)	(0.11)
REV					-0.13	-0.14	-0.14	-0.14	-0.14
					(-8.91)	(-8.89)	(-9.04)	(-9.15)	(-9.19)
ILLIQ						0.22	0.19	-0.01	-0.02
						(1.74)	(1.48)	(-0.12)	(-0.18)
IVOL							0.00	0.19	0.38
							(-0.02)	(1.94)	(4.34)
RSKEW								-0.12	-0.12
								(-13.62)	(-13.63)
RKURT								0.45	0.45
								(7.19)	(7.34)
CSK									0.26
									(0.91)
CKT									1.03
									(6.79)

Table 5: Continued.

Panel B: Dep. Var : Stock price crash probability

	1	2	3	4	5	6	7	8	9
Intercept	0.09	0.10	0.05	0.06	0.06	0.08	0.10	0.12	0.12
	(18.40)	(16.10)	(5.01)	(7.94)	(7.68)	(7.12)	(8.63)	(10.57)	(10.47)
PFM	0.64	0.70	0.32	0.36	0.36	0.31	0.29	0.54	0.53
	(1.66)	(1.80)	(0.78)	(0.88)	(0.88)	(0.77)	(0.67)	(1.27)	(1.26)
BETA		-0.96	-1.56	-1.65	-1.79	-1.89	-1.61	-1.84	-2.29
		(-3.75)	(-7.04)	(-7.69)	(-8.07)	(-10.75)	(-10.10)	(-10.65)	(-11.21)
SIZE			0.44	0.37	0.39	0.19	0.09	-0.07	-0.12
			(7.65)	(6.50)	(6.77)	(1.41)	(0.67)	(-0.52)	(-0.92)
BM				-0.69	-0.69	-0.70	-0.74	-0.80	-0.80
				(-3.10)	(-3.11)	(-3.34)	(-3.47)	(-3.86)	(-3.94)
REV					-0.02	-0.02	-0.01	-0.01	-0.01
					(-2.68)	(-2.61)	(-1.46)	(-1.21)	(-1.20)
ILLIQ						-0.14	-0.12	-0.23	-0.23
						(-1.61)	(-1.24)	(-2.55)	(-2.56)
IVOL							-0.48	-0.35	-0.22
							(-7.39)	(-6.23)	(-4.77)
RSKEW								-0.07	-0.07
								(-10.05)	(-10.14)
RKURT								0.02	0.02
								(0.40)	(0.46)
CSK									0.21
									(0.92)
CKT									0.75
									(6.16)

Table 6: Portfolio transition matrix.
This table reports the time-series mean of portfolio transition matrix. Specifically, I first sort all stocks in the sample into quintiles based on stock's past performance (PFM) calculated in month t-1 to produce losers (quintile 1) and winners (quintile 5). Then, I independently sort all stocks into quintiles by their maximum daily return (MAX) in month t. The intersections of five PFM groups and five MAX groups generate total 25 portfolios. The entry (i, j) represents that the likelihood of stocks in quintile i ranked by PFM (defined by the columns) will be in quintile j ranked by MAX (defined by the rows) in the next month. If the PFM in month t-1 and MAX in month t are completely independent, then all the probabilities should be approximately 20%. The Newey and West (1987) t statistics reported in parentheses measure the significance of transition probability higher than 20%.

	Loser	2	3	4	Winner
Low-MAX	9.91%	21.00%	26.80%	26.06%	16.31%
	(-14.25)	(2.05)	(13.62)	(11.80)	(-5.48)
2	12.75%	21.26%	24.03%	24.11%	17.82%
	(-14.22)	(5.21)	(22.56)	(14.55)	(-5.32)
3	17.27%	20.92%	20.35%	20.78%	20.66%
	(-11.15)	(6.89)	(2.77)	(5.60)	(2.28)
4	23.72%	19.77%	16.65%	16.89%	22.94%
	(11.52)	(-0.78)	(-13.26)	(-10.95)	(7.68)
High-MAX	36.36%	17.04%	12.17%	12.17%	22.28%
	(14.29)	(-7.21)	(-23.28)	(-17.91)	(2.71)

Table 7: Probability-to-win matrix.

This table reports the time-series mean of probabilities to generate an extreme high daily return in month t for each portfolio formed in month t-1. Specifically, I first construct 25 portfolios in month t-1 by independently sorting stocks based on maximum daily return (MAX) and stock's past performance (PFM) into quintile portfolios. Portfolio (1, 1) consists of stocks with low-MAX and low-PFM (loser), whereas portfolio (5, 5) is formed by stocks with high-MAX and high-PFM (winner). Next, I use MAX in month t to represent the realized extreme high return, and to sort all stocks into quintile portfolios. Finally, I calculate the proportion of stocks in each of the 25 portfolios formed in month t-1 that transformed into high-MAX portfolio in month t. The last column denoted as LMW reports the probability differences between loser and winner stocks. Newey and West (1987) t statistics are given in parentheses.

	Loser	2	3	4	Winner	LMW
Low-MAX	15.71%	5.14%	3.27%	3.33%	5.88%	9.83%
						(13.58)
2	20.16%	8.17%	5.69%	5.03%	7.83%	12.33%
						(17.57)
3	27.04%	13.70%	9.98%	8.70%	10.62%	16.42%
						(25.85)
4	37.50%	23.59%	18.94%	16.58%	17.87%	19.62%
						(34.18)
High-MAX	61.53%	46.78%	42.39%	38.32%	40.17%	21.36%
						(29.68)

Table 8: Changes in retail holdings.
This table reports the variation of retail holdings and their changes. Panel A reports the results of univariate portfolio sort based on stock's past performance (PFM). I sort all stocks in the sample into quintiles. For the stocks in each of these quintiles, I report the level (RHLD) and changes (ΔRHLD) in the number of retail holdings. I report the level (in month t-1) and changes over the months t-3 to t-1 based on the latest information during the quarter. Panels B and C report RHLD and ΔRHLD for each portfolio created by conditional bivariate portfolio sorts. Specifically, for each month, I group all stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending sort of PFM. The bottom 30% stocks are denoted as losers, and the top 30% stocks are marked as winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending sort of a lottery proxy (MAX or LTRY) calculated in month t-1. Finally, I calculate the average of retail holdings in month t-1 and their changes from month t-3 to month t-1 for each portfolio. The Newey and West (1987) t statistics are reported in parentheses.

Panel A: Univariate portfolio sort

Rank of PFM	RHLD	ΔRHLD
Loser	71.12	0.57
2	63.23	-0.18
3	61.08	-0.47
4	59.58	-0.73
Winner	64.92	-1.28
LMW	6.20	1.85
	(9.34)	(11.31)

Table 8: Continued.

Panel B: Bivariate portfolio sorts (RHLD)

	Loser	2	3	4	Winner	LMW
Low	62.02	59.63	59.46	57.58	59.44	2.58
						(3.38)
2	64.02	56.21	55.02	53.64	57.73	6.29
						(9.38)
3	68.93	59.25	56.83	54.98	61.35	7.59
						(10.41)
4	75.73	64.80	61.44	59.90	67.22	8.52
						(10.12)
High	84.89	76.26	72.63	71.80	78.84	6.05
						(6.78)
High-Low	22.87	16.63	13.17	14.23	19.40	
	(10.91)	(12.88)	(9.40)	(10.67)	(12.29)	

Panel C: Bivariate portfolio sorts (ΔRHLD)

	Loser	2	3	4	Winner	LMW
Low	0.30	-0.14	-0.30	-0.51	-0.81	1.11
						(7.61)
2	0.40	-0.25	-0.44	-0.77	-1.60	2.00
						(9.62)
3	0.55	-0.24	-0.59	-0.88	-1.67	2.22
						(10.28)
4	0.63	-0.25	-0.67	-0.96	-1.43	2.07
						(9.75)
High	1.01	-0.02	-0.36	-0.51	-0.74	1.75
						(7.79)
High-Low	0.71	0.11	-0.06	0.01	0.07	
	(4.58)	(1.18)	(-0.64)	(0.06)	(0.42)	

Table 9: Lottery premium and earnings announcements.
Panel A reports the value-weighted abnormal announcement returns for each of 25 portfolios created by conditional bivariate sorts. Specifically, in portfolio formation month t-1, I perform conditional bivariate portfolio sorts based on stock's past performance (PFM) and maximum daily return (MAX), and then calculate value-weighted portfolio abnormal returns in the next month. Abnormal announcement return are defined as the difference between the actual buy-and-hold return over the event days $t = -1$ to $t = 1$ and the expected Fama and French (1993) three-factor buy-and-hold return. In panel B, the left panel presents Fama-Macbeth regressions of abnormal announcement returns on lottery-like features measures (MAX or LTRY). The right panel presents Fama-Macbeth regressions of analyst forecast errors on the lottery-like features measures, where analyst forecast error for an announcement is the consensus forecast minus actual earnings, scaled by stock price at the end of the previous month. I control for market beta, firm size, momentum, return reversals, illiquidity, idiosyncratic volatility, realized skewness, realized kurtosis, co-skewness and co-kurtosis. The regression specification also includes interaction of LOSER, which takes a value of 1 if the stock belongs to the bottom quintile based on its past performance, and 0 otherwise, and PROXY to investigate their dependent effects on dependent variables. The Newey and West (1987) t-statistics are reported in parentheses.

Panel A: Abnormal earnings ann. ret. sorted on stock's performance and MAX

	Loser	2	3	4	Winner	LMW
Low	0.59	0.47	0.12	0.14	0.17	0.42
	(3.60)	(4.34)	(1.46)	(1.62)	(1.42)	(2.15)
2	0.30	0.73	0.22	0.21	0.37	-0.07
	(1.40)	(5.01)	(1.65)	(1.88)	(2.73)	(-0.27)
3	0.40	0.30	0.42	0.26	0.43	-0.04
	(2.01)	(1.85)	(3.40)	(1.62)	(3.05)	(-0.17)
4	0.12	0.56	0.38	0.27	0.17	-0.05
	(0.60)	(2.84)	(2.76)	(1.64)	(0.83)	(-0.18)
High	-0.58	0.30	0.46	0.33	0.00	-0.58
	(-2.75)	(1.52)	(2.15)	(1.92)	(0.01)	(-2.40)
High-Low	-1.16	-0.17	0.34	0.18	-0.17	
	(-3.92)	(-0.79)	(1.41)	(1.02)	(-0.73)	

Panel B: Fama-Macbeth regressions

	Dep. Var: Abnormal Earnings ann. ret.		Dep. Var: Analyst forecast errors	
	Proxy = MAX	Proxy = LTRY	Proxy = MAX	Proxy = LTRY
LOSER	-0.38	-0.53	0.22	0.54
	(-4.31)	(-3.15)	(1.89)	(3.87)
Proxy	0.06	0.00	1.08	0.00
	(0.04)	(-3.11)	(0.24)	(-0.63)
Proxy×LOSER	-9.32	-0.01	10.26	0.01
	(-3.07)	(-3.04)	(2.69)	(4.11)
Controls	YES	YES	YES	YES

Table 10: Short-sale constraint and optimistic beliefs.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts with the sample restriction on stocks that have options. Specifically, for each month, I group stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending sort of stock's past performance (PFM). The bottom 30% stocks are denoted as losers and the top 30% stocks are marked as winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending sort of a lottery proxy (MAX or LTRY) calculated in month $t-1$. The intersections of the three PFM groups and the five lottery proxy groups generate 15 value-weighted portfolios. Finally, I calculate portfolios' mean holding period returns in month t. Panel A reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD), and liquidity factor (PS) for each of the 15 portfolios with MAX as the lottery proxy. Panel B presents the FFCPS alphas of each portfolios with LTRY as the lottery proxy. Newey and West (1987) t statistics are given in parentheses. The sample is from 1996 to 2017.

Panel A: MAX as the lottery proxy

	Loser	Medium	Winner	LMW
Low	0.76	0.32	0.07	0.69
	(3.20)	(2.78)	(0.48)	(2.33)
2	0.47	0.14	-0.16	0.63
	(2.54)	(1.15)	(-0.88)	(2.10)
3	0.28	0.22	-0.16	0.44
	(1.08)	(1.51)	(-0.82)	(1.19)
4	-0.14	-0.09	-0.14	0.01
	(-0.44)	(-0.62)	(-0.60)	(0.02)
High	-0.30	-0.15	-0.61	0.31
	(-0.89)	(-0.79)	(-1.81)	(0.61)
High-Low	-1.06	-0.47	-0.68	-0.38
	(-2.55)	(-1.93)	(-1.68)	(-0.68)

Panel B: LTRY as the lottery proxy

	Loser	Medium	Winner	LMW
Low	0.80	0.22	-0.11	0.90
	(4.42)	(2.69)	(-0.82)	(3.42)
2	-0.11	0.03	-0.12	0.01
	(-0.60)	(0.26)	(-0.67)	(0.03)
3	-0.19	0.22	-0.26	0.07
	(-0.76)	(1.69)	(-1.57)	(0.21)
4	0.23	0.10	-0.18	0.42
	(0.81)	(0.65)	(-0.75)	(0.99)
High	0.68	0.25	-0.12	0.81
	(2.40)	(1.59)	(-0.66)	(2.15)
High-Low	-0.11	0.03	-0.02	-0.10
	(-0.41)	(0.16)	(-0.08)	(-0.26)

Table 11: Investor sentiment and lottery premium.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts during different investor sentiment periods. I use Baker and Baker and Wurgler (2004) (BW) sentiment index to separate different sentiment periods, and I define a high (low) sentiment month as one in which the BW index is above (below) the sample median value. In each sentiment period, I first perform a sequential sort by creating portfolios ranked by stock's performance (PFM) using breakpoints of 30%-40%-30%. Then, within each PFM portfolio, I form a second set of quintile portfolios ranked on stock's lottery-like features (MAX, or LTRY). Panel A reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD), and liquidity factor (PS) for each of the 15 portfolios with MAX as the lottery proxy. Panel B presents the FFCPS alphas of each portfolios with LTRY as the lottery proxy. Newey and West (1987) t statistics are given in parentheses.

Panel A: MAX as the lottery proxy

	High Sentiment		Low Sentiment	
	Loser	Winner	Loser	Winner
Low	0.79	0.03	0.22	0.23
	(5.10)	(0.25)	(1.21)	(1.71)
2	0.33	-0.27	0.01	-0.02
	(2.01)	(-2.46)	(0.03)	(-0.11)
3	-0.39	-0.29	-0.10	-0.02
	(-1.81)	(-1.64)	(-0.43)	(-0.10)
4	-1.08	-0.52	-0.32	-0.15
	(-4.32)	(-2.15)	(-1.23)	(-0.69)
High	-2.25	-0.50	-0.58	-0.29
	(-6.97)	(-2.10)	(-1.87)	(-0.88)
High-Low	-3.04	-0.53	-0.79	-0.52
	(-7.98)	(-1.99)	(-2.16)	(-1.57)

Panel B: LTRY as the lottery proxy

	High Sentiment		Low Sentiment	
	Loser	Winner	Loser	Winner
Low	0.39	-0.20	0.06	0.04
	(3.27)	(-1.83)	(0.80)	(0.28)
2	-0.18	-0.38	-0.26	0.02
	(-1.14)	(-2.66)	(-1.43)	(0.12)
3	-0.26	-0.39	-0.05	-0.06
	(-1.28)	(-2.45)	(-0.26)	(-0.28)
4	-0.78	-0.53	-0.30	-0.06
	(-3.28)	(-3.78)	(-1.53)	(-0.31)
High	-1.64	-0.55	-0.49	0.01
	(-5.95)	(-3.17)	(-1.77)	(0.05)
High-Low	-2.02	-0.35	-0.55	-0.03
	(-7.52)	(-1.69)	(-1.96)	(-0.10)

Table 12: Fama-MacBeth cross-sectional regressions.
This table reports the result of firm-level Fama-MacBeth regressions. Every month, I run cross-sectional regressions of individual stock returns on various lagged firms characteristics. LOSER is a dummy variable that takes a value of 1 if the stock belongs to the bottom quintile based on its past performance (PFM), and 0 otherwise, PROXY denotes the proxy for stock's lottery-like payoffs (MAX or LTRY). Control variables include market beta (BETA), firm size (SIZE), capital gain overhang (CGO), momentum (MOM), return reversals (REV), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), realized skewness (RSKEW), realized kurtosis (RKURT), co-skewness (CSK) and co-kurtosis (CKT). The regression specification also includes interaction of LOSER and PROXY or CGO to investigate their dependent effects on future returns. Panel A presents the regression result estimated using ordinary least squares (OLS) method and Panel B reports the results estimated using the weighted least squares (WLS) methodology following Asparouhova et al. (2013) where each observed return is weighted by one plus the observed prior return on the stock. The Newey and West (1987) t statistics of each average slop coefficients are reported in parentheses.

	Panel A: OLS				Panel B: WLS			
	Proxy = MAX		Proxy = LTRY		Proxy = MAX		Proxy = LTRY	
Intercept	1.52	4.91	0.65	4.94	1.43	5.24	0.64	5.24
	(6.59)	(7.84)	(2.76)	(8.35)	(6.15)	(8.28)	(2.65)	(8.88)
LOSER	-0.57	-0.42	-0.19	-0.17	-0.47	-0.35	-0.19	-0.19
	(-4.67)	(-4.64)	(-1.39)	(-1.78)	(-3.92)	(-3.85)	(-1.40)	(-1.82)
Proxy	0.06	-2.43	0.92	0.14	0.59	-3.72	0.97	0.19
	(0.02)	(-0.65)	(1.21)	(0.99)	(0.18)	(-0.99)	(4.29)	(1.21)
Proxy×LOSER	-18.62	-13.01	-0.66	-0.36	-17.50	-12.69	-0.75	-0.46
	(-7.13)	(-5.62)	(-4.43)	(-2.72)	(-6.76)	(-5.40)	(-4.84)	(-3.29)
Proxy×CGO		-2.90		-0.19		-3.10		-0.26
		(-1.52)		(-2.01)		(-1.92)		(-2.21)
BETA		-0.08		-0.10		-0.09		-0.11
		(-0.58)		(-0.71)		(-0.71)		(-0.80)
SIZE		-0.34		-0.36		-0.38		-0.39
		(-6.86)		(-7.58)		(-7.53)		(-8.17)
BM		0.15		0.14		0.16		0.14
		(2.31)		(2.07)		(2.42)		(2.17)
MOM		0.52		0.64		0.55		0.68
		(2.86)		(3.55)		(3.08)		(3.78)
REV		-0.03		-0.02		-0.03		-0.02
		(-7.69)		(-6.32)		(-7.11)		(-6.12)
ILLIQ		-0.13		-0.13		-0.16		-0.15
		(-4.22)		(-4.31)		(-4.94)		(-4.71)
IVOL		0.05				0.10		
		(1.26)				(2.54)		
RSKEW		-0.01		-0.01		-0.01		-0.01
		(-5.34)		(-4.49)		(-5.12)		(-4.41)
RKURT		0.08		0.02		0.08		0.03
		(4.63)		(1.22)		(4.45)		(1.33)
CSK		-0.03		-0.03		-0.01		0.00
		(-0.47)		(-0.38)		(-0.12)		(0.04)
CKT		0.12		0.10		0.11		0.10
		(4.08)		(3.42)		(4.02)		(3.45)

Table 13: Robust tests: subperiod analysis.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts during different periods. In each subperiods, I first perform a sequential sort by creating portfolios ranked by stock's past performance (PFM) using breakpoints of 30%-40%-30%. Then, within each PFM portfolio, I form a second set of quintile portfolios ranked on stock's lottery-like features (MAX). Panel A reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD), and liquidity factor (PS) for each of the 15 portfolios during the first subperiod (from 1965 to 1995). Panel B presents FFCPS alphas for each portfolios during the second period (from 1996 to 2017). Newey and West (1987) t-statistics are given in parentheses.

Panel A: 1962 - 1995

	Loser	Medium	Winner	LMW
Low	0.53	0.31	0.27	0.26
	(4.43)	(2.95)	(2.66)	(1.46)
2	0.19	0.19	-0.17	0.37
	(1.27)	(2.07)	(-1.45)	(1.67)
3	-0.37	0.02	-0.23	-0.13
	(-2.38)	(0.23)	(-1.82)	(-0.64)
4	-0.89	-0.16	-0.46	-0.43
	(-4.91)	(-1.20)	(-3.00)	(-1.82)
High	-2.12	-0.56	-0.58	-1.53
	(-9.51)	(-3.69)	(-3.33)	(-5.43)
High-Low	-2.65	-0.87	-0.86	-1.80
	(-10.34)	(-4.75)	(-4.19)	(-5.71)

Panel B: 1996 - 2017

	Loser	Medium	Winner	LMW
Low	0.63	0.35	0.08	0.54
	(3.16)	(3.55)	(0.63)	(1.95)
2	0.20	0.04	-0.14	0.34
	(0.89)	(0.45)	(-0.94)	(1.05)
3	-0.30	0.12	-0.23	-0.07
	(-1.07)	(0.97)	(-1.10)	(-0.18)
4	-0.69	-0.31	-0.41	-0.28
	(-2.33)	(-1.79)	(-1.43)	(-0.60)
High	-1.29	-0.63	-0.59	-0.70
	(-3.31)	(-3.02)	(-1.63)	(-1.50)
High-Low	-1.92	-0.98	-0.67	-1.24
	(-4.22)	(-3.57)	(-1.73)	(-2.53)

Table 14: Robust tests: price restrictions.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts with the sample restriction that stock's price is larger than \$5. For each month, I group all stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending sort of stock's past performance (PFM). The bottom 30% stocks are denoted as losers and the top 30% stocks are marked as winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending sort of a lottery proxy (MAX or LTRY) calculated in month t-1. The intersections of the three PFM groups and the five lottery proxy groups generate 15 value-weighted portfolios. Finally, I calculate portfolios' mean holding period returns in month t. Panel A reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD), and liquidity factor (PS) for each of the 15 portfolios with MAX as the lottery proxy. Panel B presents the FFCPS alphas of each portfolios with LTRY as the lottery proxy. Newey and West (1987) t statistics are given in parentheses.

Panel A: MAX as the lottery proxy

	Loser	Medium	Winner	LMW
Low	0.59	0.25	0.20	0.39
	(5.07)	(2.13)	(2.05)	(2.46)
2	0.32	0.04	-0.07	0.39
	(3.10)	(0.59)	(-0.69)	(2.33)
3	0.04	0.00	-0.15	0.19
	(0.32)	(0.07)	(-1.47)	(1.09)
4	-0.27	0.00	-0.35	0.08
	(-1.65)	(-0.04)	(-2.29)	(0.35)
High	-1.03	-0.30	-0.44	-0.59
	(-4.53)	(-1.89)	(-1.41)	(-2.77)
High-Low	-1.62	-0.54	-0.64	-0.98
	(-5.34)	(-2.03)	(-2.26)	(-3.13)

Panel B: LTRY as the lottery proxy

	Loser	Medium	Winner	LMW
Low	0.42	0.05	-0.05	0.46
	(4.13)	(1.12)	(-0.54)	(2.88)
2	0.08	0.10	-0.18	0.27
	(0.71)	(1.41)	(-1.95)	(1.47)
3	-0.12	0.10	-0.26	0.13
	(-1.01)	(1.49)	(-2.29)	(0.77)
4	0.17	0.07	-0.06	0.23
	(1.30)	(0.69)	(-0.44)	(1.14)
High	-0.80	0.00	-0.14	-0.66
	(-3.16)	(0.04)	(-1.09)	(-2.84)
High-Low	-1.22	-0.05	-0.10	-1.12
	(-4.52)	(-0.44)	(-0.64)	(-3.83)

Table 15: Robust tests: alternative measures of stock's past performance.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts with alternative measures of stock's past performance (PFM). For each month, I group all stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending sort of PFM. The bottom 30% stocks are denoted as losers and the top 30% stocks are marked as winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending sort of a lottery proxy (MAX) calculated in month $t-1$. The intersections of the three PFM groups and the five lottery proxy groups generate 15 value-weighted portfolios. Finally, I calculate portfolios' mean holding period returns in month t. Panel A reports the risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD), and liquidity factor (PS) for each of the 15 portfolios when I use 12-month cumulative return as the proxy for PFM. Panel B presents the FFCPS alphas of each portfolios when I use 6-month cumulative return as the proxy for PFM. Newey and West (1987) t statistics are given in parentheses.

Panel A: 12-month cumulative return as the proxy for PFM

	Loser	Medium	Winner	LMW
Low	0.60	0.24	0.21	0.38
	(5.87)	(3.25)	(2.56)	(3.46)
2	0.03	0.05	-0.11	0.13
	(0.24)	(0.73)	(-1.49)	(1.15)
3	-0.40	-0.09	-0.07	-0.33
	(-2.96)	(-1.24)	(-0.72)	(-2.02)
4	-0.91	-0.24	-0.27	-0.64
	(-6.40)	(-2.48)	(-2.17)	(-3.77)
High	-1.64	-0.26	-0.33	-1.31
	(-7.89)	(-2.24)	(-1.92)	(-6.01)
High-Low	-2.23	-0.51	-0.54	-1.69
	(-8.63)	(-3.59)	(-2.49)	(-5.95)

Panel B: 6-month cumulative return as the proxy for PFM

	Loser	Medium	Winner	LMW
Low	0.64	0.30	0.10	0.54
	(5.73)	(4.19)	(0.98)	(3.21)
2	0.16	0.13	-0.13	0.29
	(1.35)	(1.80)	(-1.46)	(1.76)
3	-0.17	0.07	-0.22	0.04
	(-1.14)	(0.81)	(-2.06)	(0.23)
4	-0.55	-0.15	-0.38	-0.17
	(-3.23)	(-1.46)	(-2.79)	(-0.80)
High	-1.59	-0.51	-0.47	-1.12
	(-7.91)	(-3.90)	(-2.49)	(-5.25)
High-Low	-2.23	-0.80	-0.57	-1.66
	(-8.55)	(-5.10)	(-2.50)	(-5.88)

Table 16: Robust tests: persistence of the lottery premium.
This table presents the value-weighted portfolio results for the conditional bivariate portfolio sorts for different holding periods. For each month, I group all stocks in the sample into three portfolios with 30%-40%-30% breakpoints based on an ascending sort of stock's past performance (PFM). The bottom 30% stocks are denoted as losers and the top 30% stocks are marked as winners. Within each PFM portfolio, I then sort stocks into quintile portfolios based on an ascending sort of a lottery proxy (MAX) calculated in month $t-1$. The intersections of the three PFM groups and the five lottery proxy groups generate 15 value-weighted portfolios. I report the one-month-to six-month-ahead average risk-adjusted returns (alphas) relative to the Fama-French-Carhart-Pastor-Stambaugh model (FFCPS) that includes the Fama-French three factors, momentum factor (UMD), and liquidity factor (PS) for each of the 15 portfolios. Newey and West (1987) t statistics are given in parentheses.

	1-Month		2-Month		3-Month		4-Month		6-Month	
	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner	Loser	Winner
Low	0.57	0.15	0.47	-0.06	0.37	-0.07	0.36	-0.07	0.27	-0.03
	(5.22)	(1.75)	(5.13)	(-1.05)	(4.57)	(-1.31)	(4.83)	(-1.31)	(4.15)	(-0.70)
2	0.18	-0.15	0.23	-0.15	0.13	-0.12	0.08	-0.12	0.07	-0.10
	(1.50)	(-1.57)	(2.40)	(-2.08)	(1.37)	(-2.05)	(0.95)	(-2.15)	(0.41)	(-2.15)
3	-0.23	-0.17	-0.12	-0.06	-0.18	-0.04	-0.10	-0.06	-0.02	-0.06
	(-1.45)	(-1.47)	(-1.15)	(-0.64)	(-1.84)	(-0.48)	(-1.23)	(-0.75)	(-1.98)	(-1.00)
4	-0.75	-0.43	-0.37	-0.24	-0.39	-0.25	-0.15	-0.24	-0.04	-0.20
	(-4.23)	(-2.75)	(-2.53)	(-1.84)	(-2.94)	(-1.97)	(-2.18)	(-2.19)	(-3.10)	(-1.97)
High	-1.56	-0.52	-0.96	-0.29	-0.60	-0.24	-0.24	-0.30	-0.07	-0.34
	(-6.64)	(-2.33)	(-5.38)	(-1.56)	(-4.64)	(-1.31)	(-2.53)	(-1.70)	(-0.95)	(-2.17)
High-Low	-2.13	-0.67	-1.43	-0.22	-0.97	-0.17	-0.59	-0.23	-0.34	-0.30
	(-7.73)	(-2.70)	(-7.00)	(-1.11)	(-4.98)	(-0.83)	(-2.93)	(-1.17)	(-1.26)	(-1.66)