Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Lottery-Related Anomalies: The Role of Reference-Dependent Preferences

Li An, Huijun Wang, Jian Wang, Jianfeng Yu

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 66, No. 1, January 2020, pp. 473–501
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Lottery-Related Anomalies: The Role of Reference-Dependent Preferences

Li An,[a] Huijun Wang,[b,c] Jian Wang,[d] Jianfeng Yu[a]

[a] PBC School of Finance, Tsinghua University, 100083 Beijing, China; [b] Faculty of Business and Economics, University of Melbourne, Carlton, Victoria 3010, Australia; [c] Lerner College of Business and Economics, University of Delaware, Newark, Delaware 19716; [d] School of Management and Economics, Shenzhen Finance Institute, CUHK-Business School, The Chinese University of Hong Kong, Shenzhen, China

Contact: anl@pbcsf.tsinghua.edu.cn, http://orcid.org/0000-0002-6810-5074 (LA); huijun.wang1@unimelb.edu.au (HW); jianwang@cuhk.edu.cn (JW); yujf@pbcsf.tsinghua.edu.cn, http://orcid.org/0000-0003-4309-3414 (JY)

Received: April 30, 2017
Revised: May 1, 2018; July 28, 2018
Accepted: August 20, 2018
Published Online in Articles in Advance: August 1, 2019

https://doi.org/10.1287/mnsc.2018.3205

Abstract. Previous empirical studies find that lottery-like stocks significantly underperform their non-lottery-like counterparts. Using five different measures of the lottery features in the literature, we document that the anomalies associated with these measures are state dependent: the evidence supporting these anomalies is strong and robust among stocks where investors have lost money, whereas among stocks where investors have gained profits, the evidence is either weak or even reversed. Several potential explanations for such empirical findings are examined, and we document support for the explanation based on reference-dependent preferences. Our results provide a unified framework to understand the lottery-related anomalies in the literature.

History: Accepted by Tyler Shumway, finance.

Funding: L. An and J. Yu acknowledge financial support from the National Natural Science Foundation of China [Grants 71790591 and 71790605] and Minsheng Wealth Management Research Center at PBC School of Finance, Tsinghua University.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2018.3205.

Keywords: prospect theory • lottery • reference point • skewness • default • failure probability • capital gains overhang

1. Introduction

Numerous studies have found that lottery-like stocks tend to significantly underperform nonlottery-like stocks using various measures of lottery features. A popular explanation is that investors have a strong preference for lottery-like assets, leading to the overpricing of these assets. In the data, lottery-like assets usually have a small chance of earning extremely high returns. The overweighting of the probability of these extremely high returns could, in theory, induce a strong preference for lottery-like assets (e.g., Barberis and Huang 2008). Indeed, the overweighting of small probability events is a key feature of prospect theory (PT) utility. The explanation based on the probability weighting implies an *unconditional* preference for lottery-like assets: investors prefer lottery-like assets regardless of their prior performance.[1] However, we document in this paper that the evidence for the lottery-related anomalies depends on whether investors are in a gain or loss region relative to a reference point.

Following prior studies, we use five proxies to measure the extent to which a stock exhibits lottery-like payoffs (i.e., large skewness): maximum daily returns, predicted jackpot probability, expected idiosyncratic skewness, failure probability, and bankruptcy probability. All of these measures are related to each other in that lottery-like assets under these measures exhibit large skewness in returns, although they are motivated under different concepts. Therefore, we use skewness, lottery, and lottery-like features of a stock interchangeably hereafter. We document that the relationship between the skewness and future returns is state dependent. Specifically, we first separate stocks with capital gains from those with capital losses by using the method of Grinblatt and Han (2005) to calculate the capital gains overhang (CGO) for individual stocks. CGO is essentially stock returns relative to a reference price, with positive CGO indicating capital gains relative to the reference price and vice versa. As a robustness check, we also compute an alternative measure of CGO based on the actual holdings of mutual fund managers following Frazzini (2006).

Next, we sort all individual stocks into portfolios based on lagged CGO and the five measures of lottery features in the literature. It is shown that the evidence for lottery-related anomalies is very strong and robust among stocks with capital losses (negative CGO). In contrast, the evidence for lottery-related anomalies among stocks with large capital gains (i.e., large and positive CGO) is either very weak or even reversed. For instance, we find that, among stocks with large prior capital losses (bottom quintile of CGO), the returns of lottery-like stocks (those in the top quintile of maximum daily returns in the previous month) are 138 basis points (bps) lower per month than those of nonlottery-like firms (those in the bottom quintile of maximum daily returns in the previous month). In sharp contrast, among firms with large prior capital gains (top quintile of CGO), the returns of lottery-like stocks measured by maximum daily returns are 54 bps higher per month than those of nonlottery-like stocks. Similar results hold when the lottery feature is measured by predicted jackpot probability, expected idiosyncratic skewness, failure probability, and bankruptcy probability. In addition, our results still hold when we control for a battery of additional variables, such as firm size, the book-to-market ratio, share turnover, and return volatility in the regressions of Fama and MacBeth (1973).

These findings suggest that the lottery-related anomalies depend on whether investors are in the gain or loss territory relative to a reference point. Moreover, our results are robust across all of the five lottery measures, although these measures were initially motivated by different concepts. Our empirical findings suggest that a common underlying force may have played a crucial role in all of these anomalies, and understanding these anomalies calls for a unified framework. Therefore, we go on to examine several possible explanations for our empirical findings. For the first explanation, we investigate the roles of reference-dependent preferences (RDPs) and mental accounting (MA) in these lottery-related anomalies. The key idea underlying MA is that decision makers tend to mentally frame different assets as belonging to separate accounts and then apply RDP to each account by ignoring possible interaction among these assets. The MA of Thaler (1980, 1985) provides a theoretical foundation for studies in which decision makers set a reference point for each asset they own.

With RDP, investors' risk-taking behavior in the loss region can be different from that in the gain region. For example, PT posits that individuals tend to be risk seeking in the loss region. In addition, individuals could also have a strong desire to break even after prior losses relative to a reference point (the break-even effect). Lottery-like assets are particularly attractive in these cases because they provide a better chance to recover prior losses. Thus, the current holders who are in losses are less likely to sell these lottery stocks. In other words, the effective demand for lottery stocks is particularly high when average investors of these stocks are in losses, leading to especially large overvaluation of these assets. However, when investors face prior gains, their demand for lottery-like assets is not as strong, because they are not risk seeking or in need of breaking even. Instead, because of the high volatility of lottery-like stocks, investors with MA tend to dislike these stocks if they are risk averse in their gain region.

As a result, if arbitrage forces are limited, lottery-like stocks could be overvalued compared with nonlottery-like stocks among the stocks where investors face prior losses, leading to lower future returns than nonlottery-like stocks. By contrast, among the stocks where investors face capital gains, lottery features may not be associated with lower future returns. The correlation can even turn positive because investors with capital gains usually dislike the high volatility of lottery-like stocks. Thus, RDP together with MA can potentially account for the empirical findings documented in this paper. We provide a more detailed argument in Section 3. However, we acknowledge that the static argument here might not be valid in a dynamic setting, as shown in Barberis and Xiong (2009). Although developing a formal model in a dynamic setting to account for our empirical findings would be helpful, it is beyond the scope of this paper, and therefore, we leave it for future research.

A second possible explanation for our empirical findings is from a potential underreaction to news channel as documented in Zhang (2006). To see why, we take the failure probability as an example. Stocks with capital losses (low CGO) are likely to have experienced a series of bad news. If prices respond slowly to information (underreaction to news), stocks with low CGO tend to be overvalued on average. Moreover, this underreaction effect is likely to be more severe among firms with higher failure probability, because when there is more information uncertainty (related to failure probability), investors' behavioral biases are likely to be stronger (e.g., Daniel et al. 1998, 2001) and arbitrage forces tend to be more limited. Consequently, among the stocks with low CGO, those with higher failure probabilities are likely to be more overvalued, leading to lower future returns (a negative relationship between the failure probability and future returns). However, firms with capital gains (high CGO) have probably experienced good news and therefore have been underpriced because of the underreaction to news. Similarly, this underpricing effect should be stronger for firms with higher failure probabilities, leading to higher future returns. Thus, there is a positive relationship between the failure probability and future returns among firms with high CGO. To summarize, CGO is empirically related to news experienced in the past, whereas the lottery proxy is related to information uncertainty, which is likely to exacerbate the underreaction to news effect. Therefore, the underreaction to news channel could potentially generate the empirical return pattern that we document.

The third possible explanation is from the disposition effect-induced mispricing effect. One might argue that CGO itself is a proxy for mispricing as in Grinblatt and Han (2005). Because of the disposition effect (i.e., investors' tendency to sell securities with prices that have increased since purchase rather than those with prices that have dropped), firms with higher CGO experience greater selling pressure and thus, are underpriced. Because stocks with greater skewness, especially for firms close to default, tend to have higher arbitrage costs, the final mispricing effect should be stronger among these firms. Similar to the underreaction to news story, this disposition effect-induced mispricing effect can potentially induce a negative skewness-return relation among low-CGO firms and a positive skewness-return relation among high-CGO firms as in our empirical findings. Notice that the mechanism based on RDP is different from this mispricing story because RDP does not require CGO to be a proxy for mispricing. It only needs investors' demand for skewness depending on a reference point. In addition, the lottery measures reflect return skewness in the explanation based on RDP, whereas they are proxies for arbitrage risks for the story based on the mispricing effect.

To investigate the roles of these possible mechanisms in driving our empirical findings, we perform a series of Fama and MacBeth (1973) regressions to control for (1) the interaction terms of our lottery proxies and a proxy for past news and (2) the interaction terms of the lottery proxies and a proxy for mispricing. The effect of CGO on the lottery-related anomalies remains statistically significant and quantitatively similar to that in our benchmark results. These findings suggest that our empirical results are not likely driven by CGO being a proxy for investors' underreaction to news or the mispricing (e.g., from the disposition effect). Rather, investors' high demand for lottery-like assets after prior losses may have played a critical role in our key results.

Furthermore, our main empirical findings hold up well in a variety of robustness checks. For instance, we find similar results when using different subsamples, such as excluding Nasdaq (National Association of Securities Dealers Automated Quotation) stocks or illiquid stocks. Results from the value-weighted Fama–MacBeth regressions also show that our findings are not mainly driven by small firms. In addition, the effect of CGO on the lottery-related anomalies is stronger among firms with lower institutional ownership or lower nominal stock prices, because more individuals are investing in these stocks. A similarly stronger effect is observed after high investor sentiment periods when the market participants tend to be more irrational and may be more likely to display RDP.

In the rest of this section, we relate our paper to previous studies. A large strand of literature documents that lottery-like assets have low subsequent returns. Campbell et al. (2008) show that firms with a high probability of default have abnormally low average future returns. Conrad et al. (2014) further document that firms with a high probability of default also tend to have a relatively high probability of extremely large returns (i.e., jackpot), and these firms usually earn abnormally low average future returns. Boyer et al. (2010) find that expected idiosyncratic skewness and future returns are negatively correlated. Bali et al. (2011) show that maximum daily returns in the past month are negatively associated with future returns.[2] All of these empirical studies suggest that positively skewed stocks can be overpriced and earn lower future returns. In addition, several studies have used option data to study the relation between various skewness measures and future returns of options: for instance, see Xing et al. (2010), Bali and Murray (2013), and Conrad et al. (2013).

We differ from the above studies by showing that the negative skewness–return relation is much more pronounced among firms with prior capital losses. Among firms with large prior capital gains, the empirical evidence for this negative relation is weak, insignificant, or even reversed. Our findings suggest that, in addition to an unconditional preference for skewness, such as the overweighting of small probability extreme returns, other forces also play a significant role in the lottery-related anomalies.[3] In particular, we find supportive evidence for RDP being an important source for lottery-related anomalies other than other potential explanations.

Our paper is also related to existing theoretical and empirical studies that explore the role of reference points in asset prices. Barberis and Huang (2001) find that loss aversion and MA improve a model's performance to match stock returns in the data. Barberis et al. (2001) theoretically explore the role of RDP (in particular, prospect theory) in asset prices in equilibrium settings. These studies suggest that RDP can play an important role in explaining asset pricing dynamics and cross-sectional stock returns.[4] More recently, Barberis and Xiong (2012) and Ingersoll and Jin (2013) provide theoretical models of realization utility with RDP. Our paper offers empirical support for RDP and MA, which are studied in these theoretical papers.[5]

Empirically, Grinblatt and Han (2005) find that past stock returns can predict future returns because past returns can proxy for unrealized capital gains. Frazzini (2006) shows that PT/MA induces underreaction to news, leading to return predictability. In a related study, Wang et al. (2017) show that RDP may have also played an important role in the lack of a positive risk–return trade-off in the data. Although both our study and that of Wang et al. (2017) are on the role of RDP, our paper differs from that study by focusing on the effect of RDP on lottery-related anomalies rather than the risk–return trade-off in Wang et al. (2017). In particular, we show that the effect of CGO on lottery-related anomalies is distinct from the effect of CGO on the risk–return trade-off. Therefore, our results in this paper are not primarily driven by investors' RDP for the volatility risk studied in Wang et al. (2017), although lottery-like assets tend to have higher volatility. More specifically, we use the residual skewness measures that are orthogonal to volatility, and we still find a similar effect of CGO on the residual skewness–return relation. Our paper uses similar ingredients to account for a wide range of asset pricing phenomena, which provide additional validation of the importance of the RDP channel in asset price movements. All these findings strongly suggest that the effect of RDP is pervasive rather than an artifact in the data.[6]

The rest of the paper is organized as follows. Section 2 defines the skewness proxies used in our empirical studies and presents our main findings based on these skewness proxies. Section 3 discusses several possible explanations for our empirical findings, with special attention being paid to RDP. Additional robustness tests are also reported in this section. Section 4 includes concluding remarks.

2. State-Dependent Skewness–Return Relation

This section presents our empirical finding that the skewness–return relationship depends on CGO. To proceed, we first describe our data and define the key variables used in the empirical analysis. Next, the summary statistics, double-sorting portfolio results, and Fama–MacBeth regressions results are reported.

Our data are obtained from several sources. Stock data are from the monthly and daily CRSP (Center for Research in Security Prices) database, accounting data are from the Compustat Annually and Quarterly database, and mutual fund holdings data are obtained from the Thomson Financial Mutual Funds database. We first use the data of all U.S. common stocks traded on the NYSE (New York Stock Exchange), AMEX (American Stock Exchange), and Nasdaq from 1962 to 2014 to construct various stock-level variables at the monthly frequency. After obtaining these firm–month observations, we filter our sample by requiring all observations to have nonnegative book equity, prices equal to or greater than $5, and at least 10 nonmissing daily stock returns within a month at the time of portfolio formation.

2.1. Definitions of Key Variables

This subsection describes our measures of CGO and lottery features used in previous lottery-related anomalies. More details on these key variables are provided in Online Appendix I.

2.1.1. CGO. Two CGO measures are constructed by following previous studies.

CGO^{GH}: Grinblatt and Han (2005) propose a turnover-based measure to calculate the reference price and CGO.[7] By definition, CGO is the return of a stock relative to a reference price. In Grinblatt and Han (2005), the reference price is simply a weighted average of past stock prices. The weight given to each past price is based on past turnover, which reflects the fraction of stocks that are purchased at a certain date and have not been sold since then. Therefore, the reference price is an estimate of the average purchasing price of a stock. Following Grinblatt and Han (2005), we truncate the estimation of the reference price at five years and rescale the weights to sum to one. Because we use prior five-year data to construct CGO, this CGO variable in our data ranges from January 1965 to December 2014. Moreover, a minimum of 150 weeks of nonmissing values over the past five years is required in the CGO calculation.

CGO^{FR}: In addition to the turnover-based measure of CGO, we adopt an alternative measure using mutual fund holding data as in Frazzini (2006).[8] Similar to Grinblatt and Han (2005), Frazzini (2006) defines CGO as the percentage deviation of a reference price to the current price, but this construction of reference price is arguably more accurate in capturing the average purchase price because it uses the actual net purchases by mutual fund managers. The advantage of this approach is that it can exactly identify the fraction of the shares that were purchased at a previous date and are still currently held by the original buyers. However, because of the limitation on data availability, the sample period of CGO^{FR} is shorter, ranging from April 1980 to October 2014. Also, this approach assumes that mutual fund managers are representative for all shareholders.[9]

2.1.2. Lottery Measures. We use five variables to proxy for the lottery feature of stocks following prior studies. This section briefly describes how these measures are calculated. Additional details on the construction of these measures are provided in Online Appendix I.

Maxret: Bali et al. (2011) document a significant and negative relation between the maximum daily return over the past month and the returns in the future. They also show that firms with larger maximum daily returns have higher return skewness. It is conjectured that the negative relation between the maximum daily return and future returns is caused by investors' preference for lottery-like stocks. Following their study, we use each stock's maximum daily return (*Maxret*) within the previous month as our first measure of the lottery feature.

Jackpotp: Conrad et al. (2014) show that stocks with a high predicted probability of extremely large payoffs earn abnormally low subsequent returns. Their finding suggests that investors prefer lottery-like payoffs that are positively skewed. Thus, we use the predicted probability of jackpot (*Jackpotp*; log returns greater than 100% over the next year), which is estimated from their baseline model (panel A of table 3 of Conrad et al. 2014) as our second measure. The out-of-sample predicted jackpot probabilities start in January 1972 in our paper.

Skewexp: Boyer et al. (2010) estimate a cross-sectional model of expected idiosyncratic skewness and find that it negatively predicts future returns. We use the expected idiosyncratic skewness (*Skewexp*) estimated from their model (model 6 of table 2 of Boyer et al. 2010) as our third measure. Following their estimation, this measure starts in January 1988.

Deathp: Campbell et al. (2008) find that stocks with a high predicted failure probability earn abysmally low subsequent returns. Because distressed stocks tend to have positive skewness, they conjecture that investors have a strong preference for positive skewness, which drives up the prices of distressed stocks and leads to lower future returns. We construct this proxy as our fourth measure of the lottery feature using their logit model (12-month lag in table 4 of Campbell et al. 2008). The sample period of *Deathp* starts in January 1972 because of the availability of the quarterly Compustat data used in the calculation.

Oscorep: Finally, Ohlson (1980) develops a model to predict a firm's probability of bankruptcy from a set of accounting information. He finds that firms with a higher bankruptcy probability earn lower subsequent returns. Following his approach, we calculate firms' predicted bankruptcy probability based on the O-score (*Oscorep*; model 1 of table 4 of Ohlson 1980) and use this proxy as our fifth measure of the lottery feature.

All of the five variables above are associated with return skewness in the data, although they are motivated by different concepts in the original studies.[10] We will show that they exhibit another common feature: the anomalies related to these measures depend on whether CGO is positive or negative. Then we provide a unified framework to understand all of these lottery-related anomalies.

2.2. Summary Statistics and One-Way Sorts

This section reports summary statistics and the results for single-sorted portfolios. Then Section 2.3 studies the role of CGO in the lottery-related anomalies.

Table 1 presents summary statistics and the results when stocks are sorted on lottery proxies. At the end of month t, we sort stocks into quintiles based on CGO (panel A of Table 1) or one of the five lottery proxies (panel B of Table 1). In each quintile, the portfolio excess return Ret^e is calculated as the value-weighted returns of individual stocks minus the one-month Treasury bill rate in month $t+1$. The intercepts of the Fama–French three-factor regression for the value-weighted portfolios are denoted by α_{FF3}. We also calculate other firm characteristics, such as the book-to-market value for each quintile. In these calculations, stocks are equally weighted. All firm characteristics are measured at the end of month t, with the only exception that ex post skewness is measured by return skewness over the next 12 months. All t statistics (in parentheses) are based on the heteroskedasticity-consistent standard errors of White (1980) for portfolio returns and the standard errors of Newey and West (1987) with a lag of 36 for firm characteristics.

Panel A of Table 1 reports summary statistics for portfolios sorted on CGO using the measures from both Grinblatt and Han (2005) and Frazzini (2006). Consistent with the previous literature, high-CGO firms tend to have larger firm size, higher past returns, and lower return volatility than low-CGO firms. In particular, stocks with capital gains (high CGO) outperform stocks with capital losses (low CGO) in the following month. The spread between the top and bottom quintiles is 18 bps per month. In addition, the spread between the Fama–French three-factor α values for the high- and low-CGO portfolios is 37 bps for the measure of Grinblatt and Han (2005) and 39 bps for the measure of Frazzini (2006). The spread is statistically significant for both measures. Untabulated results show that the CGO portfolio spreads tend to be more significant when January is excluded or when portfolios are equally weighted.

Panel B of Table 1 presents monthly excess returns and the Fama–French three-factor α values for portfolios sorted on the lottery proxies. Consistent with previous studies on each of these anomalies, lottery-like portfolios (row P5 of Table 1) underperform non-lottery-like portfolios (row P1 of Table 1), and the return difference is significant, especially in terms of the Fama–French three-factor α values. For instance, the Fama–French three-factor α value spread between rows P5 and P1 is 52 bps with a t-statistic of −3.74 if the lottery feature is measured by the maximum daily return in the last month. Similar results hold for other lottery proxies.

Panel B of Table 1 also reports ex post skewness for each portfolio, which is measured by the time-series mean of the cross-sectional average stock-level skewness calculated from daily stock returns in the next 12 months. As expected, we usually find that ex post skewness increases monotonically from non-lottery-like (row P1) portfolios to lottery-like (row P5) portfolios for all five lottery proxies. For instance, if the lottery feature is measured by the predicted jackpot probability, ex post skewness increases from 0.17 for row P1 to 0.60 for row P5. The difference between rows P5 and P1 is significant, and similar results hold for other lottery proxies. This result confirms that our lottery proxies, calculated at the portfolio formation time, can successfully capture stocks' lottery feature in the future. Lastly, we would like to point out that investor-perceived skewness could differ from ex post skewness because of the possible misperception of investors. Thus, the perceived differences in the skewness between lottery stocks and nonlottery stocks could be larger or smaller than the ex post skewness differences reported in panel B of Table 1.

Panel C of Table 1 reports the correlation between CGO variables, lottery proxies, and several volatility and risk measures. Total return volatility (*RetVol*) is defined as the standard deviation of monthly returns over the past five years with a minimum of two years. Idiosyncratic volatility (*IVol*) is defined as the standard deviation of the residuals from the Fama–French three-factor model using daily excess returns within a month with a minimum of 10 nonmissing observations. The variable β is CAPM (Capital Asset Pricing Model) β defined as the coefficient of the monthly CAPM regression ($R_{i,t} - R_{ft} = \alpha + \beta_{i,M}(R_{M,t} - R_{ft}) + \varepsilon_{i,t}$) over the past five years with a minimum of two years. As

Table 1. Summary Statistics

Panel A: VW excess returns and EW firm characteristics for five CGO portfolios

	Ret^e	α_{FF3}	CGO	LOGME	BM	Ret_{-1}	$Ret_{-12,-1}$	$Ret_{-36,-12}$	RetVol	Turnover
CGO of Grinblatt and Han (2005)										
CGO1	0.49	−0.14	−0.63	5.06	0.90	−0.01	−0.07	0.41	0.13	0.07
CGO2	0.41	−0.16	−0.20	5.47	0.88	0.00	0.05	0.39	0.11	0.07
CGO3	0.48	−0.04	−0.04	5.74	0.87	0.01	0.15	0.40	0.11	0.07
CGO4	0.49	0.00	0.08	5.80	0.87	0.03	0.28	0.44	0.11	0.07
CGO5	0.67	0.23	0.25	5.39	0.91	0.05	0.56	0.55	0.12	0.06
P5 − P1	0.18	0.37	0.87	0.33	0.01	0.06	0.63	0.14	−0.01	−0.01
t-Statistic	(1.01)	(2.06)	(14.77)	(1.87)	(0.22)	(12.68)	(18.76)	(2.55)	(−0.93)	(−2.09)
CGO of Frazzini (2006)										
CGO1	0.80	−0.09	−0.66	5.42	0.76	−0.03	−0.07	0.57	0.14	0.11
CGO2	0.59	−0.15	−0.15	5.81	0.77	0.00	0.08	0.45	0.12	0.08
CGO3	0.66	−0.02	0.02	6.05	0.78	0.01	0.17	0.43	0.11	0.08
CGO4	0.61	−0.02	0.15	6.20	0.77	0.03	0.30	0.47	0.11	0.08
CGO5	0.88	0.30	0.36	5.93	0.80	0.07	0.60	0.55	0.13	0.09
P5 − P1	0.08	0.39	1.01	0.51	0.04	0.10	0.67	−0.02	−0.01	−0.01
t-Statistic	(0.38)	(1.99)	(16.43)	(4.30)	(0.96)	(13.64)	(11.70)	(−0.20)	(−2.21)	(−2.56)

Panel B: VW excess returns and EW ex post skewness for five lottery portfolios

	Maxret			*Jackpotp*			*Skewexp*			*Deathp*			*Oscorep*		
Proxy	Ret^e	α_{FF3}	*ExpSkew*	Ret^e	α_{FF3}	*ExpSkew*	Ret^e	α_{FF3}	*ExpSkew*	Ret^e	α_{FF3}	*ExpSkew*	Ret^e	α_{FF3}	*ExpSkew*
P1	0.49	0.07	0.45	0.54	0.08	0.17	0.78	0.19	0.16	0.58	0.19	0.32	0.45	0.12	0.35
P2	0.52	0.03	0.38	0.68	0.05	0.30	0.65	0.01	0.22	0.55	0.05	0.28	0.55	0.10	0.37
P3	0.57	0.03	0.41	0.62	−0.07	0.40	0.65	−0.07	0.31	0.52	−0.13	0.36	0.52	−0.05	0.39
P4	0.53	−0.09	0.45	0.47	−0.28	0.49	0.24	−0.63	0.42	0.58	−0.22	0.46	0.56	−0.02	0.44
P5	0.26	−0.46	0.56	0.01	−0.75	0.60	0.14	−0.77	0.63	0.16	−0.83	0.63	0.51	−0.16	0.61
P5 − P1	−0.24	−0.52	0.11	−0.53	−0.83	0.43	−0.64	−0.96	0.46	−0.41	−1.02	0.32	0.06	−0.28	0.25
t-Statistic	(−1.07)	(−3.74)	(2.74)	(−1.74)	(−5.44)	(13.29)	(−2.06)	(−3.70)	(16.55)	(−1.64)	(−5.00)	(11.02)	(0.42)	(−2.73)	(14.80)

Panel C: Correlation matrix

	CGO^{GH}	CGO^{FR}	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*	*RetVol*	*IVol*	β
CGO^{GH}	1.00									
CGO^{FR}	0.76	1.00								
Maxret	−0.09	−0.10	1.00							
Jackpotp	−0.17	−0.24	0.54	1.00						
Skewexp	−0.18	−0.20	0.29	0.60	1.00					
Deathp	−0.38	−0.42	0.25	0.44	0.36	1.00				
Oscorep	−0.02	−0.02	0.19	0.30	0.24	0.44	1.00			
RetVol	−0.07	−0.12	0.49	0.62	0.36	0.21	0.29	1.00		
IVol	−0.18	−0.20	0.84	0.67	0.37	0.36	0.23	0.56	1.00	
β	−0.09	−0.11	0.29	0.25	0.03	0.07	0.07	0.66	0.31	1.00

Notes (to Table 1). Panel A reports the time-series averages of the monthly value-weighted (VW) excess returns (Ret^e), the intercepts of the Fama–French three-factor regression (α_{FF3}), and equal-weighted (EW) firm characteristics for five portfolios sorted by CGO. At the beginning of every month, we sort NYSE/AMEX/Nasdaq common stocks into five groups based on the quintile of the ranked values of CGO of the previous month. The portfolio is rebalanced every month. We consider two versions of CGO. The CGO of Grinblatt and Han (2005) at week t is computed as one less the ratio of the beginning of the week t reference price to the end of week $t-1$ price. The week t reference price is the average cost basis calculated as $RP_t = k^{-1}\sum_{n=1}^{T}\big(V_{t-n}\prod_{\tau=1}^{n-1}(1-V_{t-n-\tau})\big)P_{t-n}$, where V_t is week t's turnover (turnover is trading volume divided by number of shares outstanding) in the stock, T is the number of weeks in the previous five years, and k is a constant that makes the weights on past prices sum to one. Monthly CGO is weekly CGO of the last week in each month. The CGO of Frazzini (2006) at month t is defined as one less the ratio of month t reference price to the end of month t stock price. Month t reference price is an estimate of the cost basis to the representative investor as $RP_t = \phi^{-1}\sum_{n=0}^{t} V_{t,t-n}P_{t-n}$, where $V_{t,t-n}$ is the number of shares at month t that are still held by the original month $t-n$ purchasers, P_t is the stock price at the end of month t, ϕ is a normalizing constant. LOGME is the logarithm of a firm's market cap, BM is the book value of equity divided by market value at the end of the last fiscal year, Ret_{-1} is the return in the last month, $Ret_{-12,-1}$ is the cumulative return over the past year with a one-month gap, $Ret_{-36,-12}$ is the cumulative return over the past three years with a one-year gap, and *RetVol* is return volatility of the monthly returns over the past five years, Turnover is calculated as average monthly trading volume divided by number of shares outstanding over the past twelve months (where the volume is the reported value from CRSP for NYSE/AMEX stocks: 62% of CRSP reported value after 1997 and 50% of that before 1997 for Nasdaq stocks; Anderson and Dyl 2005), and *ExpSkew* is the ex post skewness calculated from daily returns over the next year. Panel B reports the time-series averages of the monthly value-weighted excess returns, the intercepts of the Fama–French three-factor regression, and equal-weighted firm ex post skewness for five portfolios sorted by each of the five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). At the beginning of every month, we sort stocks into five groups based on the quintile of the ranked values of each lottery proxy of the previous month. The portfolio is rebalanced every month. The sample period is from January 1965 to December 2014 for the CGO of Grinblatt and Han (2005), *Maxret*, and *Oscorep*; from January 1972 to December 2014 for *Jackpotp* and *Deathp*; from January 1980 to October 2014 for the CGO of Frazzini (2006); and from January 1988 to December 2014 for *Skewexp*. Panel C reports the time-series averages of the cross-sectional correlation of the relevant variables. *RetVol* is total return volatility defined as the standard deviation of monthly returns over the past five years with a minimum of two years. *IVol* is idiosyncratic volatility, defined as the standard deviation of the residuals from the Fama–French three-factor model using daily excess returns within a month with a minimum of 10 nonmissing observations. β is CAPM β defined as the coefficient of the monthly CAPM regression ($R_{i,t} - R_{ft} = \alpha + \beta_{i,M}(R_{M,t} - R_{ft}) + \varepsilon_{i,t}$) over the past five years with a minimum of two years. Monthly excess returns and Fama–French three-factor α values are reported in percentages. The t-statistics are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980) for returns and the adjusted standard errors with lag = 36 of Newey and West (1987) for firm characteristics. We always require our stocks to have nonnegative book equity, stock price equal to or greater than \$5, and at least 10 nonmissing daily stock returns in the previous month.

expected, the correlations between each pair of the lottery proxies are all positive, ranging from 0.19 to 0.6, with the two default probability measures generally having a lower correlation with the other three variables. CGO measures are generally slightly negatively correlated with lottery measures, particularly for *Deathp*, *Skewexp*, and *Jackpotp*, where past return is an explicit input in construction of the variables. Not surprisingly, stocks with higher lottery features also tend to have higher volatility. In particular, the correlation between idiosyncratic volatility and *Maxret* is 0.84, which is consistent with the findings in Bali et al. (2011). We show later that our results on lottery features remain strong and robust after controlling for volatility measures using various parametric and nonparametric approaches.

2.3. Double Sorts

As shown in the preceding subsection, our five lottery measures unconditionally predict future returns in a way that is consistent with previous studies in the literature. We now examine to what extent these predictive patterns depend on stocks' previous capital gains or losses. At the end of month t, we independently sort stocks into quintiles based on CGO and one of our five lottery measures. We next track value-weighted portfolio returns in month $t+1$.

Table 2 presents the double-sorting results based on the CGO of Grinblatt and Han (2005) and the five proxies for the lottery-like feature. Panel A of Table 2 reports excess returns for these portfolios, whereas panel B of Table 2 presents the Fama–French three-factor α values. Because of the independent sorting, we have a similar spread for the lottery proxy in the high-CGO group (CGO5) and the low-CGO group (CGO1). However, the future returns exhibit distinct patterns in these two groups. We take the maximum daily return in the last month (*Maxret*) as an example. After previous losses (CGO1), high-*Maxret* stocks *underperform* low-*Maxret* stocks by 1.38% per month in excess returns, with the t statistic equal to −5.35. In contrast, after previous gains (CGO5), the negative correlation between *Maxret* and future returns is reversed: high-*Maxret* stocks *outperform* low-*Maxret* stocks by 0.54% per month, and the t statistic is also significant at 2.30. As a comparison, the unconditional return spread between high- and low-*Maxret* portfolios is about −0.24% per month (Table 1), with the t statistic equal to −1.07. Column C5 − C1 reports the differences between lottery spreads (P5 − P1) among high-CGO firms and those among low-CGO firms. For *Maxret*, this difference-in-differences is 1.92% per month, with a t-statistic of 7.50.

The other four proxies display similar patterns. In particular, the difference-in-differences are 1.86%, 0.75%, 1.16%, and 1.15% per month for *Jackpotp*, *Skewexp*, *Deathp*, and *Oscorep*, respectively, indicating that lottery anomalies are significantly stronger among prior losers. In addition, this skewness–return pattern also

Table 2. Double-Sorted Portfolio Returns by the CGO of Grinblatt and Han (2005) and Lottery Proxies

Panel A: Excess return

	Maxret				*Jackpotp*				*Skewexp*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	1.04	0.64	0.51		0.80	0.59	0.60		0.77	0.77	0.94	
P3	0.59	0.48	0.81		0.58	0.41	1.13		0.58	0.64	0.96	
P5	−0.34	0.12	1.05		−0.37	−0.12	1.29		−0.03	−0.19	0.89	
P5 − P1	−1.38	−0.52	0.54	1.92	−1.16	−0.71	0.69	1.86	−0.80	−0.96	−0.05	0.75
t-Statistic	(−5.35)	(−2.31)	(2.30)	(7.50)	(−4.15)	(−2.16)	(2.30)	(7.36)	(−2.29)	(−2.74)	(−0.22)	(2.23)

	Deathp				*Oscorep*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	0.89	0.53	0.78		0.66	0.48	0.63	
P3	0.79	0.53	0.81		0.58	0.40	0.71	
P5	−0.04	0.57	1.02		0.04	0.40	1.16	
P5 − P1	−0.93	0.04	0.24	1.16	−0.62	−0.08	0.53	1.15
t-Statistic	(−3.04)	(0.16)	(0.85)	(3.77)	(−2.81)	(−0.48)	(2.99)	(4.70)

Panel B: FF3 α

	Maxret				*Jackpotp*				*Skewexp*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	0.52	0.16	0.10		0.22	0.08	0.19		0.02	0.13	0.40	
P3	−0.10	−0.14	0.35		−0.31	−0.33	0.55		−0.35	−0.12	0.37	
P5	−1.24	−0.60	0.45		−1.30	−0.84	0.65		−1.07	−1.07	0.16	
P5 − P1	−1.76	−0.76	0.35	2.11	−1.52	−0.92	0.46	1.98	−1.09	−1.21	−0.24	0.85
t-Statistic	(−8.36)	(−4.53)	(1.92)	(8.17)	(−7.63)	(−4.42)	(2.32)	(7.45)	(−3.59)	(−3.99)	(−1.09)	(2.52)

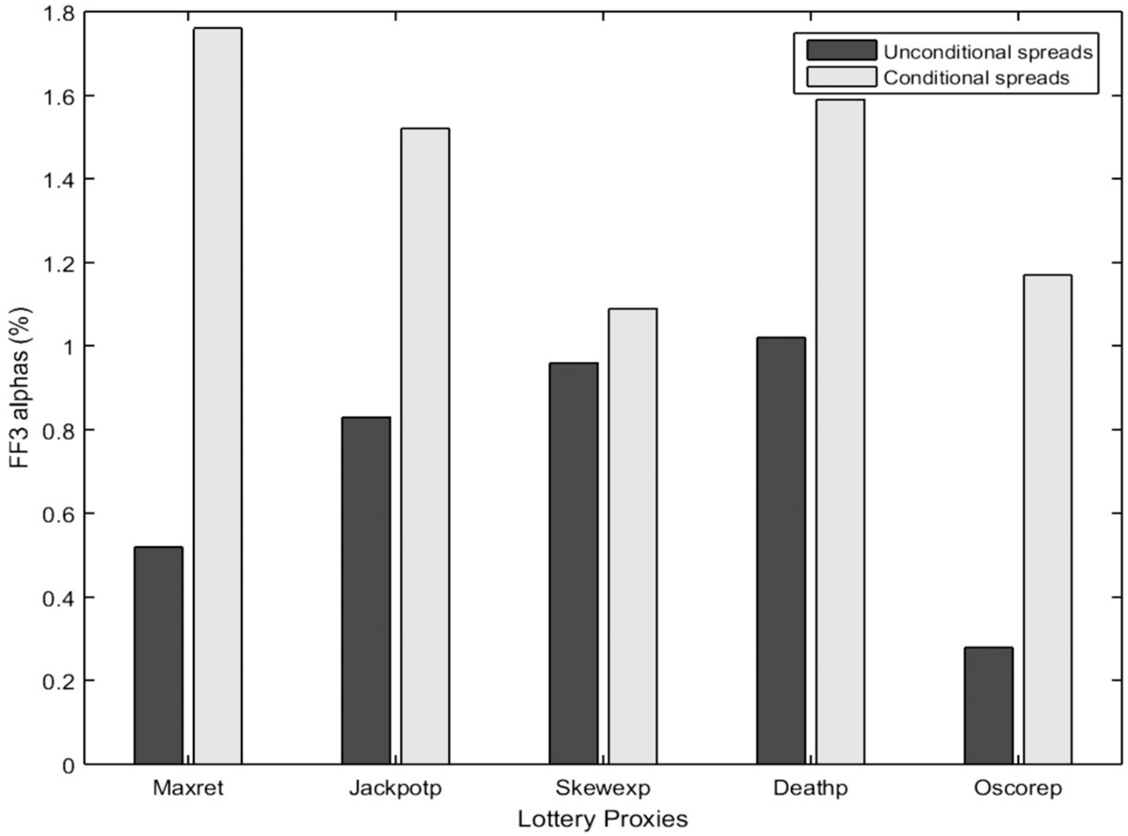
	Deathp				*Oscorep*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	0.47	0.05	0.40		0.30	0.08	0.29	
P3	0.11	−0.14	0.17		−0.16	−0.19	0.20	
P5	−1.12	−0.41	0.19		−0.87	−0.27	0.53	
P5 − P1	−1.59	−0.46	−0.21	1.38	−1.17	−0.35	0.24	1.41
t-Statistic	(−5.98)	(−1.99)	(−0.83)	(4.36)	(−6.25)	(−2.35)	(1.55)	(5.90)

Notes. At the beginning of every month, we independently sort stocks into five groups based on lagged CGO of Grinblatt and Han (2005) and five groups based on lagged lottery proxies (indicated by P1-P5). The portfolios are then held for the next month. We report the monthly value-weighted excess returns in Panel A and the intercepts of the Fama–French three-factor (FF3) regression in Panel B. The CGO of Grinblatt and Han (2005) at week *t* is computed the same way as in Table 1. Monthly CGO is weekly CGO of the last week in each month. *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability of default in the last month from Ohlson (1980). We only report the bottom-, middle-, and top-quintile CGO portfolios and their differences to save space. Excess returns and FF3 α values are reported in percentages. The sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*, from January 1972 to December 2014 for *Jackpotp* and *Deathp*, and from January 1988 to December 2014 for *Skewexp*. The *t*-statistics are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

holds for the Fama–French three-factor α values, as shown in panel B of Table 2. More interestingly, panel B of Table 2 shows that, among low-CGO firms, a large bulk of the return spreads between low- and high-skewness firms is caused by the negative α of the lottery-like assets. Taking *Maxret* as an example, the long leg has an α of 0.52% per month, whereas the short leg has an α of −1.24% per month.[11] This is consistent with the notion that facing prior losses, the demand for lottery-like assets increases. Because of limits to arbitrage and especially, short-sale impediments, this excess demand drives up the price of lottery-like assets and leads to low subsequent returns for these assets.

In contrast to low-CGO firms, the lottery-like assets do not underperform the non-lottery-like assets among high-CGO firms. In fact, among high-CGO firms, the excess return spreads between the lottery-like stocks and the non-lottery-like stocks are 0.54%, 0.69%, −0.05%, 0.24%, and 0.53% per month for the five proxies, respectively. Four of these five return spreads are positive, and three of them are significant. The patterns are similar for the Fama–French three-factor

Figure 1. Fama–French Three-Factor α Values of Unconditional and Conditional Lottery Spreads



Notes. This figure plots the time-series averages of Fama–French three-factor α spreads (in percentages) between nonlottery and lottery stocks among all of the firms and the α spreads among the firms in the bottom quintile of the CGO of Grinblatt and Han (2005). At the beginning of every month, we sort stocks into five groups based on the quintile of the ranked values of each lottery proxy of the previous month (unconditional) or independently sort stocks into five groups based on lagged CGO of Grinblatt and Han (2005) and five groups based on lagged lottery proxies (conditional). The value-weighted portfolios are then held for one month. The CGO of Grinblatt and Han (2005) at week t is computed the same way as in Table 1. Monthly CGO is weekly CGO of the last week in each month. *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability of default in the last month from Ohlson (1980). The sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*, from January 1972 to December 2014 for *Jackpotp* and *Deathp*, and from January 1988 to December 2014 for *Skewexp*. FF3, Fama–French three-factor α value.

α values, with three of five spreads being at least marginally significant and the other two negative spreads being insignificant. To provide a graphic view of the improvement on the traditional lottery strategies, Figure 1 plots the α spread between nonlottery and lottery stocks among all of the firms and the α spread among the firms with the lowest 20% CGO. It is clear that the α is increased significantly for most proxies after we constrain the universe of stocks to those low-CGO firms, where investors have especially strong demand for lottery stocks.

It is also worth noting that the lottery-like assets also underperform the non-lottery-like assets in the mid-CGO group (CGO3). These stocks are generally neither winners nor losers, with a CGO close to zero. This finding suggests that, other than the effect of investors' stronger demand for lottery-like assets after capital losses, which is emphasized in this paper, other forces, such as probability weighting, which are proposed by previous studies, should have also played an important role in the lottery-related anomalies.

To address the concern that the CGO of Grinblatt and Han (2005) is based on price–volume approximation and could be affected by high-frequency trading volume, we use the CGO of Frazzini (2006), which is based on actual holdings of mutual funds. We repeat the double-sorting exercise after replacing the CGO of Grinblatt and Han (2005) with the CGO of Frazzini (2006). The results are reported in Table 3, and they are very similar to those in Table 2. For example, panel A of Table 3 shows that the differences between excess return spreads among high-CGO firms and those among low-CGO firms (C5 – C1) are 1.88%, 1.26%, 0.56%, 1.10%, and 0.69%, respectively, per month, with corresponding t-statistics

Table 3. Double-Sorted Portfolio Returns by the CGO of Frazzini (2006) and Lottery Proxies

Panel A: Excess return

	Maxret				*Jackpotp*				*Skewexp*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	1.40	0.80	0.86		1.08	0.69	0.78		0.87	0.69	1.10	
P3	1.14	0.62	1.00		0.53	0.64	1.05		0.89	0.64	0.89	
P5	−0.24	0.18	1.10		−0.08	−0.05	0.88		0.12	0.11	0.90	
P5 − P1	−1.64	−0.62	0.24	1.88	−1.16	−0.74	0.10	1.26	−0.75	−0.58	−0.20	0.56
t-Statistic	(−4.70)	(−2.23)	(0.83)	(5.99)	(−3.45)	(−2.23)	(0.29)	(4.09)	(−1.88)	(−1.75)	(−0.70)	(1.55)

	Deathp				*Oscorep*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	1.46	0.68	0.98		1.07	0.48	1.06	
P3	1.10	0.65	0.76		0.95	0.60	0.74	
P5	0.08	0.44	0.70		0.30	0.67	0.98	
P5 − P1	−1.39	−0.23	−0.28	1.10	−0.77	0.19	−0.08	0.69
t-Statistic	(−4.06)	(−0.90)	(−0.94)	(3.10)	(−2.94)	(0.85)	(−0.40)	(2.38)

Panel B: FF3 α

	Maxret				*Jackpotp*				*Skewexp*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	0.73	0.21	0.31		0.34	0.07	0.23		0.03	0.10	0.54	
P3	0.24	−0.16	0.33		−0.46	−0.18	0.36		−0.07	−0.06	0.28	
P5	−1.39	−0.68	0.45		−1.19	−0.92	0.16		−1.10	−0.74	0.15	
P5 − P1	−2.12	−0.89	0.14	2.26	−1.54	−1.00	−0.07	1.47	−1.13	−0.84	−0.39	0.73
t-Statistic	(−7.56)	(−4.26)	(0.61)	(7.29)	(−6.59)	(−4.76)	(−0.30)	(4.36)	(−3.12)	(−3.07)	(−1.63)	(1.98)

	Deathp				*Oscorep*			
Proxy	CGO1	CGO3	CGO5	C5 − C1	CGO1	CGO3	CGO5	C5 − C1
P1	0.92	0.11	0.48		0.40	−0.10	0.58	
P3	0.31	−0.07	0.01		0.04	−0.08	0.12	
P5	−1.16	−0.57	−0.25		−0.80	−0.15	0.33	
P5 − P1	−2.08	−0.68	−0.73	1.35	−1.20	−0.04	−0.25	0.95
t-Statistic	(−7.12)	(−2.92)	(−2.36)	(3.76)	(−4.71)	(−0.19)	(−1.32)	(3.16)

Notes. At the beginning of every month, we independently sort stocks into five groups based on lagged CGO of Frazzini (2006) and five groups based on lagged lottery proxies (indicated by P1-P5). The portfolios are then held for one month. We report the monthly value-weighted excess returns in Panel A and the intercepts of the Fama–French three-factor (FF3) regression in Panel B. The CGO of Frazzini (2006) is defined the same way as in Table 1. *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). We only report the bottom-, middle-, and top-quintile CGO portfolios and their differences to save space. Excess returns and FF3 α values are reported in percentages. The sample period is from January 1980 to October 2014 for *Maxret*, *Oscorep*, *Jackpotp*, and *Deathp* and from January 1988 to October 2014 for *Skewexp*. The *t*-statistics are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

of 5.99, 4.09, 1.55, 3.10, and 2.38, respectively, for the five lottery feature proxies. The sample period in Table 3 is shorter because of the unavailability of the mutual fund holdings data for earlier dates. As a result, the *t*-statistics are slightly lower than those in Table 2. However, the economic magnitude of the spread differences remains largely the same.[12]

In panel B of Table 3, the lottery-like and non-lottery-like spreads of α values among high-CGO firms (row P5 − P1 and column CGO5) are very close to zero, and only one of them (*Deathp*) is statistically significant. In fact, among high-CGO firms, the average α spread between low- and high-skewness firms is only −26 bps (versus an average spread of −161 bps among low-CGO firms). Thus, the evidence based on the CGO of Frazzini (2006) confirms that there are virtually no return spreads between lottery-like assets and non-lottery-like assets among firms with large capital gains (high CGO).

Looking at high-CGO firms, in some specifications (particularly when excess return and the CGO of Grinblatt and Han 2005 are used), the lottery return spreads (P5 − P1) using some of our lottery proxies become positive. This positive spread among high-CGO stocks could be consistent with the standard positive risk–return relation within the gain region because lottery is to some degree related to risk. This could also be because of exposure to standard risk factors. Indeed, we find that part of the excess return spread is driven by exposure to the size factor (lottery-like stocks tend to be smaller). After controlling for exposure to Fama–French three factors, the positive spread disappears in most of the specifications. However, for jackpot probability, the α spread among high CGO firms is still significantly positive when the CGO of Grinblatt and Han (2005) is used. In Online Appendix III, we discuss this positive spread in more detail and show that it is positive mainly because lottery-like stocks typically have higher volatility and because investors tend to be risk averse and dislike volatility when they face paper gains.[13]

To ensure that the lottery characteristics spread is similar across each CGO quintile, we use independently double-sorted portfolios for our main analysis. However, because both CGO and some lottery proxies are related to past returns, one might be concerned that our independent sorts produce highly unbalanced panels. Tables IA1 and IA2 in Online Appendix II report additional summary statistics for the double-sorted portfolio characteristics. As we can see from panel A of Table IA1 and panel A of Table IA2 in Online Appendix II, the numbers of stocks in each portfolio are indeed not equal, especially for the failure probability measure. However, the smallest average number of stocks in each portfolio is still 43. In addition, one might be concerned that, among high-CGO firms that tend to have experienced high returns in the past, the spread for ex post skewness for some proxies, such as the failure probability, might be small despite our independent sorting procedure. If the difference in ex post skewness between lottery and nonlottery portfolios is smaller among high-CGO firms than among low-CGO firms, then the smaller magnitude of the return spread between lottery and nonlottery portfolios among high-CGO firms is also expected. Panel B of Table IA1 and panel B of Table IA2 in Online Appendix II address this important concern. In fact, the spreads in ex post skewness between lottery and nonlottery portfolios among high- and low-CGO firms are similar. Take the maximum daily return as an example: the difference in ex post skewness between lottery and nonlottery portfolios is 0.15 among both high- and low-CGO quintiles. A similar pattern also holds for other lottery proxies. Thus, our independent sorts indeed produce similar lottery characteristics spreads among different CGO quintiles. Additional regression analysis is also performed in subsequent sections to further address this concern.

There is one caveat in using the raw CGO measure: because CGO may correlate with other stock characteristics, in particular, past returns and shares turnover, the results in Tables 2 and 3 could be driven by effects other than the capital gains or losses that investors face. To address this concern, we sort stocks based on the residual capital gains overhang (RCGO) after controlling for other stock characteristics. To construct RCGO, we follow Frazzini (2006) by cross-sectionally regressing the raw CGO on previous 12- and 36-month returns, the previous one-year average turnover, the log of market equity at the end of the previous month, a Nasdaq dummy, an interaction term between the turnover and previous 12-month returns, and an interaction term between the turnover and the Nasdaq dummy.

Table 4 reports the Fama–French three-factor α spreads between lottery and nonlottery portfolios (P5 − P1) for low- and high-RCGO groups in the two panels on the right. To facilitate comparison, we also include lottery spreads based on raw CGO in the two panels on the left side of Table 4, which serve as a summary of the results presented in Tables 2 and 3. For each of the five lottery proxies, panel CGO^{GH} reports the lottery spreads (P5 − P1 based on the lottery proxy) among firms with low CGO (CGO1), the lottery spreads among firms with high CGO (CGO5), and the difference between these two spreads (C5 − C1). In this panel, CGO is based on the measure of Grinblatt and Han (2005). Panel CGO^{FR} presents similar results for CGO calculated from the procedure of Frazzini (2006). The two right panels report the results for RCGO under these two measures of CGO. Using the RCGO rather than the raw CGO delivers similar results that support our hypothesis as well. Taking RCGO under the procedure of Grinblatt and Han (2005), for instance, the difference between the lottery return spread among high-RCGO firms and that among low-RCGO firms is 1.13% for *Maxret* (t = 4.55), 1.10% for *Jackpotp* (t = 3.64), 0.74% for *Skewexp* (t = 2.30), 0.83% for *Deathp* (t = 2.98), and 0.53% for *Oscorep* (t = 2.24). The difference in the lottery spread between high- and low-RCGO is usually smaller than that for raw CGO. However, the difference remains significant after we use RCGO.

Because both CGO and lottery proxies are related to past returns, one might be interested in seeing the robustness checks using the Fama–French four-factor adjustment with the additional momentum factor. To save space, these results are reported in Table IA3 in Online Appendix II. Basically, these results are very similar to those based on the Fama–French three-factor adjustment in Table 2. Taking CGO under the procedure of Grinblatt and Han (2005), for instance, the difference between the lottery return spread measured by the Fama–French four-factor α among high-CGO firms and that among low-CGO firms is 1.95% for *Maxret* (t = 7.28), 1.89% for *Jackpotp* (t = 6.68), 0.81% for

Table 4. Lottery Spread and Raw CGO/RCGO: FF3 α of Lottery Spread (P5 − P1) at Different Levels of CGO

	CGO^{GH}			CGO^{FR}			$RCGO^{GH}$			$RCGO^{FR}$		
Proxy	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1	RCGO1	RCGO5	RC5 − RC1	RCGO1	RCGO5	RC5 − RC1
Maxret	−1.76	0.35	2.11	−2.12	0.14	2.26	−1.08	0.05	1.13	−1.31	−0.15	1.16
	(−8.36)	(1.92)	(8.17)	(−7.56)	(0.61)	(7.29)	(−4.61)	(0.26)	(4.55)	(−4.66)	(−0.65)	(3.75)
Jackpotp	−1.52	0.46	1.98	−1.54	−0.07	1.47	−1.26	−0.16	1.10	−1.23	−0.56	0.68
	(−7.63)	(2.32)	(7.45)	(−6.59)	(−0.30)	(4.36)	(−6.12)	(−0.64)	(3.64)	(−4.70)	(−2.30)	(1.89)
Skewexp	−1.09	−0.24	0.85	−1.13	−0.39	0.73	−1.06	−0.32	0.74	−1.11	−0.29	0.82
	(−3.59)	(−1.09)	(2.52)	(−3.12)	(−1.63)	(1.98)	(−3.49)	(−1.16)	(2.30)	(−3.12)	(−1.10)	(2.22)
Deathp	−1.59	−0.21	1.38	−2.08	−0.73	1.35	−1.32	−0.49	0.83	−1.50	−0.70	0.81
	(−5.98)	(−0.83)	(4.36)	(−7.12)	(−2.36)	(3.76)	(−4.83)	(−2.16)	(2.98)	(−4.51)	(−2.76)	(2.29)
O-score	−1.17	0.24	1.41	−1.20	−0.25	0.95	−0.60	−0.07	0.53	−0.80	−0.31	0.49
	(−6.25)	(1.55)	(5.90)	(−4.71)	(−1.32)	(3.16)	(−3.08)	(−0.42)	(2.24)	(−3.04)	(−1.54)	(1.52)

Notes. This table reports the Fama–French three-factor (FF3) α values for the lottery spread (difference between top- and bottom-quintile lottery portfolios) of the bottom- and top-quintile CGO portfolios and their difference. Twenty-five portfolios are constructed at the end of every month from independent sorts by each one of the four CGO definitions and each one of five lottery proxies. The four CGO definitions include the CGO of Grinblatt and Han (2005) (CGO^{GH}), the CGO of Frazzini (2006) (CGO^{FR}), RCGO, and $RCGO^{GH}$ and $RCGO^{FR}$ corresponding to CGO^{GH} and CGO^{FR}, respectively. RCGO is the residual obtained by regressing cross-sectionally the raw CGO on previous 12- and 36-month returns, the previous 12-month average turnover, the log of market equity at the end of the previous month, an interaction term between turnover and previous 12-month return, and an interaction term between turnover and Nasdaq dummy. The portfolio is then held for one month. We consider five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). FF3 α values are reported in percentages. In the cases of CGO^{GH} and residual CGO^{GH}, the sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*, from January 1972 to December 2014 for *Jackpotp* and *Deathp*, and from January 1988 to December 2014 for *Skewexp*. In the cases of CGO^{FR} and residual CGO^{FR}, the sample period is from January 1980 to October 2014 for *Maxret*, *Oscorep*, *Jackpotp*, and *Deathp* and from January 1988 to October 2014 for *Skewexp*. The *t*-statistics are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Skewexp ($t = 2.25$), 1.17% for *Deathp* ($t = 3.47$), and 1.15% for *Oscorep* ($t = 4.78$). In addition, our double-sorting results are robust to equal-weighted returns. In our benchmark analysis, we focus on value-weighted portfolio returns and exclude penny firms from our sample. This approach helps to keep our results from being dominated by the behavior of very small firms, as warned by Fama and French (2008). However, the properties of value-weighted returns could be dominated by the behavior of a few very large firms because of the well-known heavy-tail distribution of firm sizes in the U.S. stock market (Zipf 1949). To address this concern, Table 5 reports the results for two alternative weighting methods: equal-weighted and lagged gross return-weighted portfolio α values.[14] The lagged gross return-weighted portfolio returns are also considered because this weighting scheme is designed to mitigate the liquidity bias in asset pricing tests (Asparouhova et al. 2013).

The results in Table 5 confirm a significant role for CGO in the lottery-related anomalies. That is, among low-CGO firms, the lottery spreads are negative and highly significant, whereas among high-CGO firms, all the lottery spreads are either positive or insignificantly negative except for the predicted failure probability (*Deathp*). The sizes of the differences in the lottery spread (C5 − C1) are very close for equal-weighted and lagged gross return-weighted portfolio returns. They are also very similar to the value-weighted portfolio returns in our benchmark results, suggesting that our findings are not mainly driven by extremely large or small firms.

In panel (III) of Table 5, we show that our results are also robust to conditional sorting. We double-sort portfolios independently in our benchmark analysis. In contrast, conditional sorting first ranks stocks based on lagged CGO. Next, we sort stocks within each CGO group according to one of the five lottery proxies. Then the value-weighted return of each portfolio is calculated in the same way as in our benchmark analysis. Panel (III) of Table 5 shows that our benchmark findings hold both qualitatively and quantitatively under conditional sorting. The differences in lottery spreads between high- and low-CGO groups (C5 − C1) are statistically significant and quantitatively similar to those in Table 2. In all panels of Table 5, the results are based on the CGO measure of Grinblatt and Han (2005). The results based on the measure of Frazzini (2006) are quantitatively similar and are not reported to save space.

2.4. Fama–MacBeth Regressions

The double-sorting approach in the preceding section is simple and intuitive, but it cannot explicitly control for other variables that may influence returns. However, sorting on three or more variables is impractical. Thus, to examine other possible mechanisms, we perform a series of Fama and MacBeth (1973) cross-sectional regressions, which allow us to conveniently control for additional variables.

In all the Fama–MacBeth regressions below, we control for a list of traditional return predictors, such as firm size, book-to-market ratio, past returns, stock return

Table 5. Equal-Weighted and Lagged Gross Return-Weighted Portfolios and Conditional Sorts

	(I) Equal weighted			(II) Lag return weighted			(III) Conditional sort		
Proxy	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1
Maxret	−1.81	0.08	1.88	−1.88	0.09	1.97	−1.74	0.25	1.99
	(−13.86)	(0.57)	(10.78)	(−14.70)	(0.64)	(11.20)	(−8.00)	(1.36)	(7.76)
Jackpotp	−1.12	0.63	1.74	−1.27	0.60	1.88	−1.72	0.34	2.06
	(−7.37)	(4.09)	(9.45)	(−8.63)	(3.72)	(10.02)	(−8.10)	(1.81)	(7.19)
Skewexp	−0.72	0.28	1.00	−0.86	0.24	1.10	−1.14	−0.28	0.86
	(−3.36)	(1.67)	(4.70)	(−4.07)	(1.36)	(5.21)	(−4.24)	(−1.13)	(2.70)
Deathp	−1.17	−0.45	0.73	−1.26	−0.51	0.75	−1.98	−0.44	1.54
	(−7.85)	(−2.78)	(3.80)	(−8.43)	(−2.95)	(3.79)	(−7.10)	(−2.30)	(4.72)
Oscorep	−0.83	0.23	1.06	−0.82	0.24	1.07	−1.24	0.30	1.54
	(−7.43)	(2.11)	(7.37)	(−7.34)	(2.18)	(7.34)	(−6.30)	(1.95)	(6.32)

Notes. This table reports the Fama–French three-factor monthly α values (in percentages) for the lottery spread (difference between top- and bottom-quintile lottery portfolios) among the bottom- and top-quintile CGO portfolios and their differences for five double-sorted robustness tests. The 25 portfolios are constructed at the end of every month from independent sorts by the CGO of Grinblatt and Han (2005) and each one of five lottery proxies in tests (I) and (II). The equal-weighted and lagged gross return-weighted portfolio α values are reported in panels (I) and (II), respectively. In panel (III), 25 portfolios are constructed from conditional sorts by first dividing stocks into five groups based on lagged CGO and further dividing stocks within each of the CGO groups into five groups based on lagged lottery proxies. The portfolio is then held for one month. The CGO of Grinblatt and Han (2005) at week t is computed as one less the ratio of the beginning of the week t reference price to the end of week $t-1$ price. The week t reference price is the average cost basis calculated as $RP_t = k^{-1}\sum_{n=1}^{T}\big(V_{t-n}\prod_{\tau=1}^{n-1}(1-V_{t-n-\tau})\big)P_{t-n}$, where V_t is week t's turnover in the stock, T is the number of weeks in the previous five years, and k is a constant that makes the weights on past prices sum to one. Turnover is calculated as trading volume divided by number of shares outstanding. Monthly CGO is weekly CGO of the last week in each month. We consider five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). The sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*. The t-statistics in parentheses are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

volatility, and share turnover. Following Conrad et al. (2014), independent variables are winsorized at their 5th and 95th percentiles. The benchmark regression in column (0) of Table 6 shows that the coefficient of CGO is significant and positive, suggesting that stocks with more unrealized capital gains have higher future returns, which confirms the finding of Grinblatt and Han (2005). Grinblatt and Han (2005) attribute this finding to investors' tendency to sell stocks with capital gains (high CGO). This overselling makes high-CGO stocks undervalued and predicts high future returns for these stocks.

Next, we investigate the role of CGO in the lottery anomalies. In Table 6, regressions in column (1) under the five lottery proxies are our main results in this section. We will discuss the results in columns (2)–(4) in the next section. Under each lottery proxy, the regressions in column (1) have two more independent variables than the benchmark regression in column (0): the lottery proxy and an interaction term between the proxy and CGO. For all five lottery proxies, the coefficient estimate of the interaction term is always positive and significant. It suggests that lottery-like stocks with negative CGO have lower returns than lottery-like stocks with positive CGO, confirming that our results based on double sorts still hold even after we control for size, book-to-market ratios, past returns, stock return volatility, and shares turnover. It is noteworthy that the coefficient of lottery proxy itself typically seems to be negative and significant, suggesting that lottery-like assets have lower future returns than non-lottery-like assets, especially when CGO is negative.

In sum, our results generally confirm the previous findings of a negative skewness–return relation in the lottery-related anomalies. However, both our portfolio and regression results highlight the role of CGO in understanding these lottery-related anomalies.

3. Possible Explanations

In this section, we compare three possible explanations for our documented dependence of the lottery-related anomalies on CGO. If the lottery proxies appropriately capture the lottery features of stocks and CGO reflects investors' status of capital gains or losses, RDP is naturally a potential explanation for our empirical findings: investors' demand for lottery-like stocks is stronger when they are in capital loss. However, if the lottery proxies mainly capture investors' speed at incorporating past news rather than stocks' lottery features, the underreaction to news documented in Zhang (2006) can also potentially account for our empirical findings. In addition, if CGO is mainly an indicator of mispricing because of the disposition effect rather than investors' status of gains or losses, our empirical results can be potentially caused by the mispricing effect too. In this section, we discuss and compare these three potential explanations in detail.

Table 6. Fama–MacBeth Regressions Using the CGO of Grinblatt and Han (2005)

	Benchmark	*Proxy = Maxret*				*Proxy = Jackpotp*			
	(0)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
CGO	0.004	−0.013	−0.015	−0.014	−0.015	−0.009	−0.008	−0.008	−0.008
	(4.07)	(−9.42)	(−10.71)	(−8.29)	(−9.06)	(−5.10)	(−5.34)	(−4.01)	(−4.02)
Proxy		0.010	0.026	−0.007	−0.001	−0.287	−0.269	−0.491	−0.471
		(0.78)	(1.97)	(−0.41)	(−0.07)	(−4.26)	(−3.75)	(−5.31)	(−4.99)
Proxy × *CGO*		0.284	0.322	0.301	0.323	1.122	1.145	1.011	0.991
		(12.84)	(13.15)	(10.56)	(11.15)	(8.69)	(8.28)	(5.92)	(5.81)
Proxy × $Ret_{-12,-2}$			−0.054		−0.059		0.075		0.062
			(−2.35)		(−2.34)		(0.64)		(0.52)
Proxy × *VNSP*				0.197	0.272			1.304	1.294
				(2.44)	(3.11)			(2.99)	(2.9)
Ret_{-1}	−0.060	−0.060	−0.060	−0.063	−0.064	−0.051	−0.051	−0.054	−0.054
	(−15.25)	(−14.49)	(−14.61)	(−15.18)	(−15.31)	(−12.33)	(−12.44)	(−12.98)	(−13.05)
$Ret_{-12,-2}$	0.009	0.009	0.012	0.007	0.010	0.008	0.007	0.006	0.005
	(6.46)	(6.33)	(7.11)	(4.86)	(5.94)	(5.37)	(3.57)	(4.11)	(2.65)
$Ret_{-36,-13}$	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002	−0.002
	(−1.61)	(−1.35)	(−1.43)	(−1.97)	(−2.00)	(−1.99)	(−2.02)	(−2.37)	(−2.44)
LOGME	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002	−0.001
	(−3.42)	(−3.13)	(−3.10)	(−3.02)	(−3.01)	(−4.60)	(−4.50)	(−4.83)	(−4.73)
LOGBM	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(2.32)	(2.33)	(2.33)	(2.65)	(2.65)	(2.31)	(2.32)	(2.54)	(2.55)
VNSP				0.008	0.004			0.004	0.004
				(1.68)	(0.86)			(0.77)	(0.78)
IVol	−0.213	−0.174	−0.181	−0.195	−0.203	−0.108	−0.111	−0.125	−0.127
	(−7.53)	(−4.59)	(−4.77)	(−5.27)	(−5.48)	(−3.82)	(−3.93)	(−4.57)	(−4.65)
β	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(2.55)	(2.62)	(2.54)	(2.39)	(2.34)	(1.67)	(1.67)	(1.58)	(1.57)
Turnover	−0.029	−0.029	−0.028	−0.030	−0.028	−0.021	−0.020	−0.021	−0.020
	(−1.91)	(−1.96)	(−1.87)	(−1.98)	(−1.88)	(−1.48)	(−1.43)	(−1.47)	(−1.43)

Table 6. (Continued)

	Proxy = Skewexp				Proxy = Deathp				Proxy = Oscorep			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
CGO	−0.007	−0.005	−0.010	−0.008	−0.001	−0.002	−0.003	−0.003	0.000	0.000	0.000	0.000
	(−3.09)	(−2.38)	(−3.61)	(−3.08)	(−0.53)	(−1.08)	(−1.98)	(−1.98)	(0.05)	(0.10)	(0.30)	(0.36)
Proxy	−0.002	−0.002	−0.004	−0.004	−10.074	−9.743	−9.892	−9.744	−0.010	−0.002	−0.050	−0.042
	(−1.42)	(−1.85)	(−2.25)	(−2.27)	(−8.45)	(−7.99)	(−6.08)	(−5.90)	(−0.82)	(−0.14)	(−2.54)	(−2.10)
Proxy × *CGO*	0.016	0.013	0.017	0.014	6.197	8.609	7.957	8.549	0.195	0.203	0.156	0.158
	(6.65)	(5.16)	(5.85)	(4.92)	(2.00)	(2.69)	(2.93)	(2.93)	(7.03)	(5.88)	(4.74)	(4.18)
Proxy × $Ret_{-12,-2}$		0.005		0.006		−2.784		0.160		0.012		−0.002
		(2.19)		(2.50)		(−1.16)		(0.06)		(0.30)		(−0.04)
Proxy × *VNSP*			0.006	0.003			0.710	1.376			0.283	0.315
			(0.90)	(0.40)			(0.08)	(0.15)			(2.29)	(2.42)
Ret_{-1}	−0.034	−0.034	−0.037	−0.038	−0.046	−0.047	−0.065	−0.065	−0.057	−0.057	−0.061	−0.061
	(−7.15)	(−7.19)	(−7.72)	(−7.77)	(−14.96)	(−15.06)	(−16.03)	(−16.09)	(−14.68)	(−14.76)	(−15.41)	(−15.49)
$Ret_{-12,-2}$	0.008	0.005	0.006	0.003	0.007	0.009	0.005	0.005	0.009	0.009	0.007	0.007
	(4.17)	(2.23)	(3.23)	(1.22)	(5.15)	(5.28)	(3.33)	(2.84)	(6.08)	(5.76)	(4.6)	(4.39)
$Ret_{-36,-13}$	−0.002	−0.001	−0.002	−0.002	−0.001	−0.001	−0.002	−0.002	−0.001	−0.001	−0.002	−0.002
	(−2.20)	(−2.07)	(−2.44)	(−2.34)	(−2.94)	(−2.96)	(−3.31)	(−3.32)	(−2.41)	(−2.39)	(−3.06)	(−3.04)
LOGME	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
	(−2.18)	(−2.25)	(−2.14)	(−2.17)	(−3.38)	(−3.36)	(−2.92)	(−2.89)	(−3.85)	(−3.79)	(−3.67)	(−3.61)
LOGBM	0.000	0.000	0.000	0.000	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001
	(−0.01)	(−0.01)	(0.18)	(0.18)	(3.85)	(3.89)	(4.32)	(4.35)	(1.67)	(1.70)	(1.98)	(2.01)
VNSP			0.015	0.017			0.021	0.022			0.019	0.019
			(2.36)	(2.70)			(4.64)	(4.79)			(5.52)	(5.39)
IVol	−0.163	−0.162	−0.181	−0.180	−0.108	−0.106	−0.131	−0.130	−0.193	−0.194	−0.212	−0.214
	(−5.19)	(−5.18)	(−6.03)	(−6.02)	(−3.84)	(−3.79)	(−4.93)	(−4.86)	(−6.61)	(−6.63)	(−7.53)	(−7.55)
β	0.003	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001
	(2.25)	(2.26)	(2.15)	(2.13)	(1.69)	(1.69)	(1.50)	(1.51)	(1.83)	(1.82)	(1.46)	(1.47)
Turnover	−0.019	−0.020	−0.014	−0.014	−0.017	−0.017	−0.015	−0.014	−0.027	−0.026	−0.026	−0.025
	(−1.74)	(−1.76)	(−1.25)	(−1.25)	(−1.17)	(−1.18)	(−1.04)	(−1.00)	(−1.77)	(−1.73)	(−1.71)	(−1.65)

Notes. Every month, we run a cross-sectional regression of returns on lagged variables. The time-series average of the regression coefficients is reported. CGO is defined as in Grinblatt and Han (2005). *LOGBM* is the log of book to equity, *LOGME* is the log of market equity, Ret_{-1} is return in the last month, $Ret_{-12,-1}$ is the cumulative return over the past year with a one-month gap, $Ret_{-36,-12}$ is the cumulative return over the past three years with a one-year gap, Turnover is average monthly turnover (i.e., monthly trading volume divided by number of shares outstanding) over the past twelve months, *IVol* is idiosyncratic volatility defined as the standard deviation of the residuals from the Fama–French three-factor model using daily excess returns within a month with a minimum of 10 nonmissing observations, and β is CAPM β, defined as the coefficient of the monthly CAPM regression ($R_{i,t} - R_{ft} = \alpha + \beta_{i,M}(R_{M,t} - R_{ft}) + \varepsilon_{i,t}$) over the past five years with a minimum of two years. *VNSP* is a measure of the V-shaped disposition effect calculated based on An (2016). *Maxret* is the maximum daily return over the past month, *Jackpotp* is the predicted jackpot probability from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness from Boyer et al. (2010), *Deathp* is the predicted failure probability from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability from Ohlson (1980). Independent variables are winsorized at their 5th and 95th percentiles. The sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*, from January 1972 to December 2014 for *Jackpotp* and *Deathp*, and from January 1988 to December 2014 for *Skewexp*. The *t*-statistics in parentheses are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

3.1. The Role of RDP

Investors are uniformly risk averse in most standard asset pricing models because these models use the expected utility function that is globally concave. This assumption has been a basic premise in numerous studies that help to understand observed consumption and investment behaviors in finance and economics.

However, RDP has recently attracted massive attention in several research fields following the seminal work by Kahneman and Tversky (1979). The idea of reference points is a critical element in the prospect theory developed by Kahneman and Tversky (1979). Their theory predicts that most individuals have an S-shaped value function, which is concave in the gain domain but convex in the loss domain. Both gains and losses are measured relative to a reference point. In addition, investors are loss averse in the sense that the disutility from losses is much higher than the utility from the same amount of gains.[15] Finally, the mental accounting of Thaler (1980, 1985) provides a theoretical foundation for decision makers setting a separate reference point for each asset that they own by ignoring possible interactions among those assets.

Building on the RDP model by Kahneman and Tversky (1979) and MA, a large number of recent studies have shown that RDP can better capture human behaviors in many decision-making processes and can account for many asset pricing phenomena that contradict the prediction of standard models.[16] Moreover, psychological and evolutionary foundations for RDP are also documented in Frederick and Loewenstein (1999) and Rayo and Becker (2007).

Among studies suggesting that investors' preferences are reference dependent, a strand of literature (e.g., Odean 1998, Grinblatt and Keloharju 2001, Dhar and Zhou 2006) finds that individual investors are averse to loss realization. Similar evidence is also found for professional investors: for instance, see Locke and Mann (2000) for a study on futures traders, Shapira and Venezia (2001) for a study on professional traders in Israel, Wermers (2003) and Frazzini (2006) for studies on mutual fund managers, and Coval and Shumway (2005) for a study on professional market makers at the Chicago Board of Trade. Although these studies focus on investors' trading behaviors as implied by RDP, our paper differs from them by investigating the asset pricing implications of RDP. In particular, we focus on cross-sectional stock return predictability as implied by investors' RDP.

Under the assumption of the reference point being the lagged status quo, the aversion to loss realization predicts investors' willingness to take unfavorable risks to regain the status quo. A related concept, the break-even effect coined by Thaler and Johnson (1990), also suggests that, after losses, investors often have a strong urge to make up their losses because by breaking even, investors can avoid having to prove that their first judgment was wrong. The break-even effect can induce investors in losses to take gambles that they otherwise would not have taken. In this case, assets with high skewness seem especially attractive because they provide a better chance to break even.

In contrast, among stocks with prior capital gains, there are two countervailing forces. On the one hand, investors might still prefer lottery-like stocks, probably because of the overweighting of small-probability events in the standard probability weighting scheme of prospect theory, although the demand for lottery-like assets becomes weaker as the effects from breaking even and aversion to loss realization disappear. Thus, the lottery-like stocks can still be moderately overvalued. On the other hand, the lottery-like stocks typically have higher (idiosyncratic) volatility. When facing prior gains, investors are risk averse and dislike even stock-level idiosyncratic volatility because of MA. Thus, the lottery-like stocks can be undervalued and exhibit high future returns. Overall, it is not clear which force dominates in the data. However, we can at least conclude from the above discussions that investors' demand for lottery-like stocks should be stronger in the loss region than in the gain region.

Below, we would like to further clarify how CGO can affect asset prices and especially how CGO can interact with lottery features in affecting asset prices. Let us start with the model in Grinblatt and Han (2005), which shows that the disposition effect can affect the equilibrium price and result in return predictability. In their model, the disposition effect at the current time point leads to a demand perturbation caused by the purchases made in previous periods. The current equilibrium price is shown to be a linear combination of the asset's fundamental value and the purchase price of the average investor; the latter part is the over- or undervaluation relative to the right price. In their model, the firms in losses (i.e., negative CGOs) are relatively overpriced but not because investors are buying those assets. These firms are overpriced because their current holders are not willing to sell their existing shares owing to the disposition effect. Effectively, there is excess demand from the current shareholders for these stocks with average investors that are in losses. This is the key insight from Grinblatt and Han (2005), and the same mechanism has also been used by Frazzini (2006).

Now consider the case of the valuation of lottery stocks. In a similar vein, the overvaluation of these assets can come from the excess demand of their current holders. For some reason, if the price at which their current holders are willing to sell is higher than the fundamental value of the lottery stock, the stock can be overvalued. The overvaluation does not have to take the form of actual purchases or sales. We propose that

RDP (for lottery) and MA can jointly explain the return patterns that we empirically documented. Specifically, when an investor faces a larger prior loss in an asset, he or she tends to have a higher (irrational) valuation for the asset's lottery feature (the same asset that he or she has the loss in because of MA) probably because such a feature provides a better chance to break even, as mentioned earlier. In other words, compared with lottery assets with average investors that are in gains, the lottery assets with average investors that are in losses face effectively a higher demand from their current holders. Thus, in a representative agent model with limits to arbitrage (or in a model like that of Grinblatt and Han 2005, in which parts of the agents are fully rational and the rest of the agents have a behavioral bias), this behavioral tendency has the following pricing implication: the overvaluation (at time t) of lottery assets relative to nonlottery assets is higher among the stocks with average investors that are in losses (at time t) than among the stocks with average investors that are in gains (at time t). Because CGO measures the average unrealized capital gains for all investors at the portfolio formation time, the return spreads between nonlottery stocks and lottery stocks (from t to $t+1$) should be higher among the firms with low CGOs than among the firms with high CGOs. Thus, CGO can interact with lottery features in affecting asset prices.

In sum, a natural implication from RDP and MA is that the lottery-related anomalies should be weaker or even reversed among stocks where investors have experienced gains, especially large gains. In contrast, the negative relationship between skewness and expected returns should be much more pronounced among stocks where investors have experienced losses and have been seeking break-even opportunities.[17]

The results in Section 2 indeed show such a pattern: a strong negative correlation between expected (abnormal) returns and skewness exists among firms with a low (negative) CGO, whereas a weak (insignificant or even reversed) correlation between expected abnormal returns and skewness exists among firms with a high (positive) CGO. Furthermore, the return spreads (between high- and low-skewness stocks) are significantly more negative among firms with capital losses than those among firms with capital gains. In addition, to better support this potential explanation, we provide disaggregated evidence on investors' trading behavior using trading data for both retail investors and mutual fund managers. Specifically, using the five skewness proxies and the same brokerage data set as in Barber and Odean (2000), we show that individual investors' demand for lottery-like assets over non-lottery-like assets is significantly stronger in the loss region than in the gain region.[18] Using probit regressions, we estimate the propensity to sell lottery-like stocks for individual investors. The coefficients for the interaction terms between unrealized returns and skewness proxies are significant, implying that individual investors exhibit a stronger demand for lottery-like assets after losses than after gains. Additionally, using mutual fund holding data, we find that mutual fund managers exhibit the same trading behavior. These results confirm our conjecture about the role of RDP in the lottery anomalies, and we discuss them in more detail in Section 3.5.

We now discuss the relation between RDP and some other popular explanations in the literature for the documented lottery-related anomalies. The overweighting of small-probability events in prospect theory can lead to the overpricing of positively skewed assets, which can potentially account for the anomalies related to maximum daily returns, predicted jackpot probability, and expected idiosyncratic skewness. In fact, our double-sorts exercises show that the lottery-related anomalies are generally significant in the middle-CGO groups, indicating a significant role of this kind of probability weighting in the lottery-related anomalies. Also, the larger default option values of distressed firms combined with shareholder expropriation could lead to the low returns of the distressed firms because the default option is a hedge (e.g., Garlappi et al. 2008, Garlappi and Yan 2011).[19]

However, the key difference between RDP and the above previous mechanisms is the heterogeneity of the lottery effect across stocks. RDP implies that the lottery-related anomalies should be much more pronounced among firms with low CGO, whereas the previous mechanisms typically predict that the anomalies should be homogeneous across different CGO levels. For example, if investors overweight small-probability events, the overweighting effect should be similar across different levels of CGO, and thus, the lottery effect should not depend on CGO.

Again, we would like to emphasize that the mechanism of RDP does not depend on the probability weighting: even without the overweighting of small-probability events, the break-even effect and the investor's desire to avoid losses could still lead to excess demand for positive skewness when investors face prior losses. Thus, RDP is distinct from the mechanisms based on probability weighting, which is the prevalent explanation for the lottery-related anomalies in the existing literature (e.g., Barberis and Huang 2008, Bali et al. 2011, Conrad et al. 2014). Our empirical findings suggest that RDP may have played a crucial role in accounting for the lottery-related anomalies, although other mechanisms are likely to work simultaneously in investors' decision-making process, and the probability weighting would be significantly amplified by the excess demand for lottery-type assets among prior losers.

Lastly, one could argue that the return spread between nonlottery and lottery firms should be negatively related to the aggregate level of CGO. However, this time-series variation in the lottery effect is not a very robust prediction of RDP because of other potential countervailing and confounding effects. Countercyclical risk aversion, for instance, predicts that investors would have relatively stronger demand for risk (including default risk) in expansions, and high aggregate CGO tends to coincide with economic booms. If firm-level risk cannot be fully diversified away, countercyclical risk aversion also predicts the opposite time-series variation in the skewness–return relation. More importantly, in aggregate, after favorable shocks (i.e., during booms), many investors may have realized profits, although the unrealized profits are also likely to be high.[20] Then, because of the standard house money effect, investors could (in aggregate) prefer high-volatility or lottery-like stocks even more, the opposite of our prediction. Notice that the house money effect does not contradict our CGO effect on the lottery return spread in the cross section because those who have realized profits are not the owners of this particular stock anymore, although they may own other stocks.

In principle, we could try to control all the time-series effects and isolate the effect of aggregate CGO on the time-series variation of the lottery spread. However, we see at least two difficulties with this approach. First, it is hard to control all possible time-series effects. That is, we may leave out some important effects that we are unaware of. Second, many of these potential effects (such as aggregate risk aversion and the house money effect) are hard to measure. This is exactly why we focus on the cross-sectional heterogeneity of the lottery return spread; that is, we mainly use a difference-in-differences approach in the cross section. In this way, our analysis is more immune to various potentially opposing time-series effects on lottery demand. In Table IA4 in Online Appendix II, we show that after controlling for some potential confounding effects, aggregate CGO indeed marginally predicts the lottery return spread with the expected sign.

3.2. Underreaction to News

Our empirical findings may also reflect that lottery-like assets react to news more slowly than non-lottery-like assets. Zhang (2006) argues that information travels slowly, which can lead to significant underreaction of asset prices to past news. This underreaction effect might be stronger among firms with higher information uncertainty, where investors' biases are likely to be stronger (e.g., Daniel et al. 1998, 2001) and arbitrage forces tend to be more limited. Thus, among the firms with recent bad news, higher information uncertainty is likely to forecast lower future returns because of the current underreaction to the past bad news.

Our proxies for the lottery-like feature could be related to information uncertainty, especially for the failure probability of Campbell et al. (2008) and the bankruptcy probability of Ohlson (1980), because these firms might indeed be hard to evaluate. Because high-CGO firms are likely to have experienced good news in the past, if lottery-like firms have high information uncertainty, a positive relation between the lottery proxies and future returns will exist in the data among high-CGO firms. Conversely, firms with low CGO are likely to have experienced negative news and have been overpriced because of news underreaction. This overpricing effect is more pronounced for lottery-like stocks because of higher information uncertainty, implying a negative relation between the lottery proxies and future returns among firms with low CGO. This argument is consistent with the skewness–return CGO pattern observed in Tables 2 and 3, and it also implies a positive coefficient for the interaction term between CGO and skewness proxies in Fama–MacBeth regressions.

To examine the importance of this underreaction to news effect in driving our empirical results, we include in the Fama–MacBeth regressions an interaction term between a proxy for the past news and our lottery proxies. Following Zhang (2006), past realized returns (the cumulative return over the past year with a one-month lag) are used as a proxy for news.[21] Regression (2) in Table 6 shows that the interaction terms of past returns and our proxies for the lottery feature ($Proxy \times Ret_{-12,-1}$) are insignificant for all the skewness proxies except for the maximum daily return of the last month and the expected idiosyncratic skewness. However, the sign of the interaction term is negative for the maximum daily return of the last month, which argues against the underreaction to news effect being an explanation for our findings. In addition, after controlling for the underreaction to news effect, the interaction terms of CGO and the lottery proxies remain significant with similar t-statistics. The t-statistics for the interaction term are 13.19 for maximum daily return, 8.22 for predicted jackpot probability, 5.39 for expected idiosyncratic skewness, 2.26 for failure probability, and 6.05 for bankruptcy probability.

3.3. CGO as a Proxy for Disposition Effect–Induced Mispricing

Other than being a proxy for aggregate capital gains or losses, CGO may also be directly related to disposition effect–induced mispricing, which could drive our empirical findings. As documented by Grinblatt and Han (2005), firms with higher CGO tend to experience higher selling pressures because of the disposition effect (investors being more likely to sell a security after a gain rather than a loss), which, in turn, leads to lower current prices and higher future returns. In general, the final mispricing effect that survived after arbitrage tends to be stronger for firms with higher limits to arbitrage. If our proxies for the lottery-like feature are related to limits to arbitrage, the positive relation between CGO and future returns can be amplified when firms have high skewness. Indeed, one may expect that firms close to default should impose higher arbitrage risk for arbitrageurs.[22] Note that, in this interpretation, the roles for CGO and lottery are reversed compared with the RDP interpretation: the RDP interpretation posits that lottery proxy is the source of mispricing and that CGO plays a moderating role by capturing investors' lottery preference–related to prior capital gains; in this interpretation, CGO itself is a proxy for mispricing, and the lottery measures are the moderating factors because they are related to limits to arbitrage. Both interpretations would lead to a positive coefficient for the interaction term between CGO and skewness proxies in Fama–MacBeth regressions, as we have documented.

To address this concern, we control for a more precise disposition effect–induced mispricing measure (relative to CGO) that is derived from the V-shaped disposition effect following An (2016). The V-shaped disposition effect is a refined version of the disposition effect: Ben-David and Hirshleifer (2012) find that investors are more likely to sell a security when the magnitude of their gains or losses on this security increases and their selling schedule, characterized by a V shape, has a steeper slope in the gain region than in the loss region. Motivated by this more precise description of investor behavior, An (2016) shows that stocks with large unrealized gains and losses tend to outperform stocks with moderate unrealized gains and losses. More importantly, the V-shaped net selling propensity (VNSP), a more precise measure of mispricing, subsumes the return predictive power of CGO.

In regression (3) of Table 6, VNSP and its interaction term with our skewness proxies are added to the Fama–MacBeth regression. The coefficient estimate of $Proxy \times VNSP$ is significant only for three of the five lottery proxies. It suggests that the mispricing effect may have played a role in some of the lottery anomalies but not all of them. More importantly, our empirical findings are not driven by the mispricing effect. After controlling for this effect, the coefficients of $Proxy \times CGO$ remain similar in magnitude to those in regression (1), and the t-statistics are positive and significant in all cases. In regression (4), we include all of the control variables in previous regressions, and the estimated coefficients of $Proxy \times CGO$ only change marginally in magnitude and remain statistically significant for all lottery proxies.

In sum, both the underreaction to news effect and the mispricing story cannot account for the skewness–return pattern that we have documented in Table 2. Coupled with the investors' trading behaviors documented in Tables 11 and 12, we believe that the stronger demand for lottery-like assets after prior losses plays a critical role in the lottery-related anomalies.

3.4. Additional Robustness Checks

We now conduct a series of additional tests to assess the robustness of our results. In the first set of results reported in Table 7, we address the following two concerns. First, one potential concern about our Fama–MacBeth regression results is that all stocks are treated equally. The standard cross-sectional regression places the same weight on a very large firm as on a small firm. Thus, the results based on equal-weighted regressions could be disproportionately affected by small firms, which account for a relatively small portion of the total market capitalization. Although the results based on equal-weighted regressions reflect the effect of a typical firm, they might not appropriately measure the effect of an average dollar. To alleviate this size effect, we perform the value-weighted Fama–MacBeth regressions in which returns are weighted by firms' market capitalizations at the end of the previous month using the same model as in column (4) in Table 6.

Second, another concern is that our empirical findings could be driven by Nasdaq or illiquid stocks. Previous studies (e.g., Bali et al. 2005, Bali and Cakici 2008) show that some asset pricing phenomena disappear after the most illiquid stocks are excluded from the sample. Thus, to address this concern, we consider a subset of stocks that can be classified as the top 90% liquid stock. Following Amihud (2002), we measure illiquidity by the average ratio of the daily absolute return to the daily dollar trading volume over the past year.

Specifically, we repeat the Fama–Macbeth regressions as in column (4) in Table 6, but now we use the following alternative specifications: (1) We use the weighted least square (WLS) regressions, where the weight equals each firm's market capitalization at the end of the previous month. (2) We exclude all Nasdaq stocks and only include stocks listed on the NYSE and the AMEX. (3) We exclude the most illiquid stocks—those that fall in the top illiquid decile in each month (using the illiquidity measure of Amihud 2002). Table 7 presents the results for these three groups of regressions. Both the coefficients and t-statistics of the interaction term between CGO and the lottery feature proxies are similar to those obtained in the Fama–MacBeth regressions of Table 6, with all the t-statistics remaining statistically significant at the 5% level. In addition, Table IA5 in Online Appendix II reports the lottery spreads of the double-sorting portfolios after excluding Nasdaq firms or illiquid firms. The results remain largely the same as in the benchmark portfolio results.

In sum, the evidence in Table 7 shows that the role of RDP in the skewness–return relationship is not driven by highly illiquid stocks or Nasdaq stocks or disproportionately affected by small firms because both the statistical significance and the economic magnitude remain largely the same after controlling for these factors.

Table 7. Fama–MacBeth Regressions, Robustness Checks

	(I) WLS					(II) Excluding Nasdaq stocks					(III) Excluding top illiquid decile				
Proxy	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*
CGO	−0.014	−0.006	−0.009	−0.008	−0.005	−0.017	−0.008	−0.010	−0.006	−0.002	−0.015	−0.008	−0.010	−0.007	−0.001
	(−7.13)	(−3.13)	(−3.11)	(−4.45)	(−3.43)	(−7.50)	(−2.96)	(−2.76)	(−2.49)	(−1.13)	(−8.07)	(−3.55)	(−3.41)	(−3.67)	(−1.03)
Proxy	0.059	−0.636	−0.003	−4.904	−0.104	0.000	−0.488	−0.001	−8.719	−0.025	0.008	−0.791	−0.006	−11.983	−0.087
	(2.73)	(−3.73)	(−0.99)	(−1.78)	(−2.89)	(−0.01)	(−3.83)	(−0.51)	(−4.39)	(−1.12)	(0.43)	(−6.43)	(−3.29)	(−6.36)	(−3.29)
Proxy × *CGO*	0.240	0.706	0.013	13.051	0.173	0.373	1.150	0.015	13.713	0.259	0.301	1.064	0.016	14.224	0.167
	(6.21)	(2.84)	(2.87)	(2.70)	(3.12)	(8.24)	(3.53)	(2.97)	(2.53)	(4.50)	(8.84)	(4.48)	(3.85)	(3.73)	(3.25)
Proxy × $Ret_{-12,-2}$	0.004	−0.010	−0.006	−1.075	0.005	−0.010	0.324	0.005	6.498	0.003	−0.048	−0.090	0.002	−1.334	−0.003
	(0.13)	(−0.05)	(−1.79)	(−0.25)	(0.09)	(−0.30)	(1.78)	(1.39)	(1.91)	(0.06)	(−1.81)	(−0.61)	(0.68)	(−0.43)	(−0.06)
Proxy × *VNSP*	0.316	3.161	0.004	7.466	0.552	0.187	1.464	0.002	−2.416	0.280	0.462	3.545	0.019	28.849	0.588
	(2.82)	(4.82)	(0.32)	(0.52)	(3.04)	(1.56)	(2.06)	(0.14)	(−0.19)	(1.75)	(4.42)	(5.97)	(1.85)	(2.41)	(3.71)
Ret_{-1}	−0.056	−0.038	−0.022	−0.043	−0.046	−0.055	−0.047	−0.023	−0.054	−0.056	−0.063	−0.053	−0.034	−0.060	−0.058
	(−10.71)	(−7.29)	(−3.60)	(−8.84)	(−9.62)	(−12.09)	(−10.24)	(−4.06)	(−12.27)	(−13.00)	(−14.24)	(−12.07)	(−6.59)	(−14.51)	(−14.21)
$Ret_{-12,-2}$	0.009	0.007	0.009	0.007	0.008	0.008	0.004	0.005	0.003	0.007	0.010	0.006	0.005	0.006	0.007
	(3.48)	(2.96)	(3.24)	(3.14)	(4.00)	(4.18)	(1.78)	(1.75)	(1.30)	(3.76)	(5.29)	(2.96)	(2.28)	(3.02)	(4.25)
$Ret_{-36,-13}$	0.000	0.000	0.000	−0.001	−0.001	0.000	−0.001	0.000	−0.001	−0.001	−0.001	−0.001	−0.002	−0.002	−0.002
	(−0.23)	(−0.35)	(−0.13)	(−0.99)	(−0.66)	(−0.58)	(−0.65)	(−0.25)	(−1.64)	(−1.19)	(−1.77)	(−2.14)	(−2.18)	(−3.21)	(−2.66)
LOGME	−0.001	−0.002	−0.001	−0.001	−0.002	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002	−0.001	−0.001	−0.001
	(−4.29)	(−4.18)	(−2.72)	(−3.55)	(−4.53)	(−3.05)	(−4.21)	(−1.55)	(−2.71)	(−3.22)	(−3.26)	(−5.26)	(−2.17)	(−3.03)	(−3.78)
LOGBM	0.000	0.001	0.000	0.002	0.000	0.001	0.002	0.000	0.003	0.001	0.001	0.001	0.000	0.002	0.001
	(0.73)	(1.54)	(−0.52)	(2.28)	(0.65)	(2.85)	(3.12)	(0.44)	(4.86)	(2.71)	(2.09)	(2.15)	(−0.21)	(3.82)	(1.62)
VNSP	−0.007	−0.009	−0.004	0.004	0.003	0.002	−0.006	−0.007	0.012	0.005	−0.002	−0.008	0.008	0.015	0.017
	(−1.18)	(−1.72)	(−0.54)	(0.75)	(0.72)	(0.26)	(−0.93)	(−0.77)	(2.01)	(1.05)	(−0.41)	(−0.46)	(1.23)	(2.84)	(4.36)
IVol	−0.409	−0.171	−0.203	−0.161	−0.200	−0.214	−0.153	−0.201	−0.149	−0.212	−0.309	−0.166	−0.218	−0.165	−0.244
	(−8.00)	(−3.97)	(−4.37)	(−3.93)	(−5.03)	(−4.8)	(−4.19)	(−4.25)	(−4.35)	(−6.17)	(−7.90)	(−5.20)	(−5.92)	(−5.21)	(−7.73)
β	0.000	0.000	0.002	0.000	0.000	0.002	0.002	0.002	0.002	0.001	0.002	0.002	0.003	0.002	0.001
	(0.40)	(0.09)	(1.21)	(−0.06)	(−0.20)	(2.17)	(1.52)	(1.76)	(1.60)	(1.14)	(2.11)	(1.42)	(2.13)	(1.36)	(1.24)
Turnover	−0.004	0.001	0.009	0.003	0.004	−0.028	−0.021	−0.014	−0.013	−0.027	−0.028	−0.021	−0.009	−0.017	−0.025
	(−0.22)	(0.08)	(0.79)	(0.18)	(0.19)	(−1.91)	(−1.49)	(−1.35)	(−0.95)	(−1.79)	(−1.86)	(−1.52)	(−0.90)	(−1.27)	(−1.65)

Notes. This table reports the time-series average of the regression coefficients from three Fama–MacBeth regressions robustness tests. In test (I), every month, we run a cross-sectional weighted least squares regression of returns on lagged variables with market equity of the last month as the weighting. In tests (II) and (III), every month, we run a cross-sectional regression of returns on lagged variables on two subsamples: Nasdaq stocks are excluded in test (II), and the top-decile illiquid stocks are excluded (using the illiquidity measure of Amihud 2002) in test (III). Variable definitions and sample period are the same as in Table 6. *LOGBM* is the log of book to equity, and *LOGME* is the log of market equity. The intercept of the regression is not reported. The *t*-statistics are in parentheses, and they are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Next, we confirm that our results are not mainly driven by investors' reference-dependent preference for return volatility. Because high-skewness stocks are typically also more volatile, it is possible that the underperformance of lottery-like assets among firms with negative CGO is caused by investors' preference for volatility (rather than skewness) after losses. For example, prospect theory posits that investors are risk-seeking after losses, and thus, they might prefer stocks with high volatility after losses. Indeed, Wang et al. (2017) find a significant and negative risk–return relation among low-CGO stocks where investors face losses. To ensure that our results are not primarily being driven by investors' preference for volatility after losses, we reexamine the patterns on lottery portfolios by purging the confounding effect from volatility. We use both parametric and nonparametric methods to control for the volatility effect, and the results are shown in Table 8.

Panels A and B of Table 8 report double-sorted portfolio results based on CGO and residual lottery measures. In particular, at each month, we first run cross-sectional regressions of each of our five lottery proxies on monthly return volatility over the past five years, and then we use the residual lottery proxies to repeat our double-sorting exercises. Panel A of Table 8 reports results using *IVol*, and panel B of Table 8 reports results using *RetVol*. In panels C and D of Table 8, we do volatility-adjusted lottery sorts to further control for the potential nonlinear relation between volatility and lottery proxies. Specifically, we first sort all stocks into 10 deciles based on *IVol* (panel C of Table 8) or *RetVol* (panel D of Table 8); within each decile, we then divide stocks into five groups based on each one of the five lottery proxies, and finally, we collapse across the volatility groups. In this way, we obtain five volatility-adjusted lottery portfolios, and each portfolio contains stocks with a similar level of volatility. We then do double-sorting exercises based on CGO and volatility-adjusted lottery. The results indicate that the pattern of the lottery spreads holds reasonably well. In particular, the differences in lottery spreads (after controlling for the volatility effect) among high- and low-CGO firms are positively significant for most of the specifications.

Moreover, we also perform Fama–MacBeth regressions to control for the interaction effect between CGO, volatility, and other variables, and the results are presented in Table 9. After adding the interaction term between CGO and volatility (using idiosyncratic volatility as a proxy in panel A of Table 9 and total return volatility in panel B of Table 9) to the regressions, the coefficients of $Proxy \times CGO$ are still strongly significant for 9 of the 10 specifications, confirming that investors' reference-dependent preference for volatility does not seem to be a main driver for our results. In other words, the evidence based on both the portfolio approach and the Fama–MacBeth regressions is consistent with the notion that stocks with higher skewness are more appealing to investors facing losses because the stocks have a better chance of breaking even.

In another robustness check, we separate the whole sample into several subsamples based on the quartiles of institutional holdings or nominal stock prices. We find that the effect of CGO on the lottery-related anomalies is generally stronger among firms with lower institutional holdings or firms with low nominal prices. These results are presented in Table 10. In particular, among firms with low institutional ownership, the differences in the lottery spread between high-CGO firms and low-CGO firms are 2.21%, 1.97%, 2.17%, 0.88%, and 1.90% for the five skewness proxies, respectively. By contrast, among firms with high institutional ownership, the differences in the lottery spread between high- and low-CGO firms are only 1.15%, 0.57%, 0.15%, 0.38%, and 0.50% for the five skewness proxies, respectively. The differences between these numbers are economically significant. The stronger effect among firms with lower institutional holdings is consistent with the limits to arbitrage effect (e.g., Nagel et al. 2005). Moreover, previous studies find a positive relationship between stock prices and institutional ownership, suggesting that individual investors prefer low-price stocks (e.g., Falkenstein 1996, Gompers and Metrick 2001, Kumar 2009, Kumar et al. 2011).[23] Thus, our evidence based on both institutional ownership and nominal prices is also consistent with the notion that the effect of the reference point on the lottery-related anomalies should be stronger among firms with more individual investors because the reference-dependent preference might be a better description of individuals' risk attitudes rather than institutional investors' risk attitudes.[24]

We then examine how the return patterns that we document vary with investor sentiment in Baker and Wurgler (2006). Stambaugh et al. (2012) show that many anomalies are stronger after high-sentiment periods when more noisy traders participate in the market. Indeed, Table IA6 in Online Appendix II confirms that the negative lottery spread among low-CGO firms is much more significant after a high-sentiment period than it is after a low-sentiment period. In addition, the role of RDP in the lottery-related anomalies is more significant after a high-sentiment period than it is after a low-sentiment period. Indeed, during high-sentiment periods, the differences in the lottery spread between high- and low-CGO firms are 2.50%, 2.64%, 0.02%, 2.30%, and 1.95% for the five skewness proxies, respectively. By contrast, during low-sentiment periods, the differences in the lottery spread between high- and low-CGO firms are only 0.56%, 0.64%, –0.63%, –0.47%, and 0.97% for the five skewness proxies, respectively. The differences between these numbers are also economically significant.

Lastly, we rerun our full empirical model using Fama–MacBeth regression (model (4) in Table 6) and

Table 8. Residual Lottery Spreads

		Maxret			*Jackpotp*			*Skewexp*			*Deathp*			*Oscorep*		
Proxy		CGO1	CGO5	C5 – C1	CGO1	CGO5	C5 – C1	CGO1	CGO5	C5 – C1	CGO1	CGO5	C5 – C1	CGO1	CGO5	C5 – C1
Panel A: Residual lottery spread, *IVol*																
Exret	P5 – P1	–0.79	0.12	0.92	–0.63	0.70	1.33	–0.40	–0.06	0.34	–0.49	0.14	0.63	–0.11	0.18	0.29
	t-Statistic	(–4.03)	(0.64)	(3.86)	(–2.57)	(2.92)	(5.41)	(–1.28)	(–0.23)	(1.02)	(–1.78)	(0.60)	(2.09)	(–0.55)	(1.01)	(1.20)
FF3 α	P5 – P1	–0.87	0.13	1.00	–0.91	0.51	1.42	–0.60	–0.24	0.36	–1.00	–0.24	0.76	–0.42	0.00	0.42
	t statistic	(–4.20)	(0.70)	(4.03)	(–4.57)	(2.77)	(5.34)	(–2.08)	(–1.04)	(1.03)	(–3.62)	(–1.1)	(2.36)	(–2.14)	(0.00)	(1.73)
Panel B: Residual lottery spread, *RetVol*																
Exret	P5 – P1	–0.92	0.10	1.02	–0.68	0.54	1.21	–0.68	–0.05	0.63	–0.60	0.17	0.77	0.05	0.02	–0.03
	t-Statistic	(–4.59)	(0.63)	(4.19)	(–2.99)	(2.62)	(4.81)	(–1.89)	(–0.18)	(1.82)	(–1.97)	(0.72)	(2.47)	(0.21)	(0.11)	(–0.11)
FF3 α	P5 – P1	–1.02	–0.02	0.99	–0.91	0.28	1.19	–0.80	–0.17	0.63	–1.19	–0.19	1.00	–0.20	–0.15	0.05
	t-Statistic	(–4.80)	(–0.14)	(3.96)	(–4.15)	(1.60)	(4.62)	(–2.57)	(–0.66)	(2.07)	(–4.07)	(–0.82)	(3.18)	(–0.97)	(–0.84)	(0.19)
Panel C: Volatility-adjusted lottery spread, *IVol*																
Exret	P5 – P1	–0.74	–0.21	0.54	–0.20	0.41	0.61	–0.44	0.03	0.47	–0.38	–0.17	0.21	–0.37	0.32	0.69
	t-Statistic	(–5.11)	(–1.66)	(3.02)	(–1.15)	(2.35)	(3.52)	(–1.67)	(0.12)	(1.63)	(–1.94)	(–0.70)	(0.84)	(–1.77)	(1.69)	(3.32)
FF3 α	P5 – P1	–0.77	–0.15	0.62	–0.38	0.18	0.56	–0.49	–0.06	0.43	–0.83	–0.47	0.36	–0.75	0.10	0.85
	t-Statistic	(–5.49)	(–1.19)	(3.34)	(–2.95)	(1.21)	(3.00)	(–1.96)	(–0.34)	(1.73)	(–4.86)	(–2.17)	(1.43)	(–3.75)	(0.58)	(3.79)
Panel D: Volatility-adjusted lottery spread, *RetVol*																
Exret	P5 – P1	–1.52	–0.08	1.44	–0.40	0.27	0.67	–0.62	0.10	0.72	–0.65	–0.09	0.56	–0.08	0.28	0.36
	t-Statistic	(–10.51)	(–0.69)	(7.88)	(–2.18)	(1.47)	(3.42)	(–2.45)	(0.46)	(2.77)	(–2.97)	(–0.41)	(2.56)	(–0.45)	(1.78)	(2.09)
FF3 α	P5 – P1	–1.55	–0.18	1.37	–0.55	0.09	0.64	–0.53	0.13	0.66	–1.10	–0.40	0.70	–0.44	0.13	0.56
	t-Statistic	(–10.33)	(–1.32)	(7.32)	(–3.27)	(0.55)	(3.46)	(–2.15)	(0.72)	(2.68)	(–5.29)	(–2.13)	(3.34)	(–3.04)	(0.87)	(3.23)

Notes. Panels A and B report the value-weighted excess returns and Fama–French three-factor (FF3) α values for residual lottery spreads (high minus low portfolio returns based on residual lottery measures) in the bottom- and top-quintile CGO groups and their differences. Twenty-five portfolios are constructed at the end of every month from independent sorts by the CGO of Grinblatt and Han (2005) and each of the five residual lottery proxies, which are the residuals obtained by regressing cross-sectionally each of the five lottery proxies on total return volatility (panel A) or idiosyncratic volatility (panel B). Panels C and D report the value-weighted excess returns and FF3 α values for lottery spreads based on volatility-adjusted lottery portfolios in the bottom- and top-quintile CGO groups and their differences. Specifically. we first sort all stocks into 10 deciles according to their total return volatility (panel C) or idiosyncratic volatility (panel D); within each decile, we then equally divide stocks into five groups according to each one of the five lottery proxies; and finally, we collapse across the volatility groups. In this way, we obtain five volatility-adjusted lottery portfolios; we interact them independently with five CGO quintiles and result in 25 portfolios. These portfolios are held for one month. *RetVol* is total return volatility, defined as the standard deviation of monthly returns over the past five years with a minimum of two years. *IVol* is idiosyncratic volatility, defined as the standard deviation of the residuals from the FF3 model using daily excess returns within a month with a minimum of 10 nonmissing observations. We consider five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). Returns and α values are reported in percentages. The sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*, from January 1972 to December 2014 for *Jackpotp* and *Deathp*, and from January 1988 to December 2014 for *Skewexp*. The *t*-statistics are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Table 9. Fama–MacBeth Regressions Controlling for the Interaction Between Volatility and CGO

	Maxret	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*
Proxy	Panel A: *IVol*					Panel B: *RetVol*				
CGO	−0.016	−0.013	−0.016	−0.022	−0.016	−0.020	−0.015	−0.016	−0.022	−0.016
	(−8.98)	(−8.05)	(−6.56)	(−10.59)	(−8.76)	(−9.22)	(−6.45)	(−5.47)	(−9.05)	(−7.23)
Proxy	0.004	−0.420	−0.004	−10.532	−0.045	0.006	−0.426	−0.003	−9.058	−0.040
	(0.25)	(−5.59)	(−3.16)	(−6.30)	(−2.22)	(0.33)	(−4.59)	(−2.11)	(−5.35)	(−2.05)
$Proxy \times CGO$	0.297	0.326	0.006	7.883	0.102	0.290	0.615	0.009	12.736	0.106
	(7.15)	(2.42)	(2.86)	(2.19)	(2.71)	(9.61)	(3.26)	(2.98)	(3.64)	(2.75)
$Proxy \times Ret_{12,-2}$	−0.059	−0.020	0.004	−1.841	0.002	−0.057	0.088	0.006	−0.828	0.006
	(−2.37)	(−0.26)	(2.61)	(−0.77)	(0.05)	(−2.30)	(0.74)	(2.60)	(−0.35)	(0.13)
$Proxy \times VNSP$	0.268	0.898	0.001	4.304	0.289	0.269	1.328	0.006	−0.904	0.302
	(3.08)	(2.96)	(0.22)	(0.52)	(2.23)	(3.10)	(2.96)	(0.72)	(−0.11)	(2.35)
$Vol \times CGO$	0.118	0.519	0.477	0.875	0.756	0.057	0.095	0.089	0.142	0.141
	(1.04)	(8.71)	(7.78)	(12.00)	(10.06)	(3.00)	(4.56)	(3.83)	(7.81)	(7.16)
Ret_{-1}	−0.064	−0.048	−0.033	−0.052	−0.063	−0.064	−0.055	−0.039	−0.050	−0.062
	(−15.37)	(−13.21)	(−7.99)	(−16.37)	(−16.15)	(−15.69)	(−13.31)	(−8.24)	(−16.03)	(−16.03)
$Ret_{-12,-2}$	0.010	0.004	0.002	0.005	0.006	0.010	0.005	0.003	0.005	0.006
	(5.91)	(2.74)	(0.99)	(3.41)	(3.88)	(5.80)	(2.56)	(1.07)	(3.16)	(3.97)
$Ret_{-36,-13}$	−0.001	−0.001	−0.001	−0.001	−0.002	−0.001	−0.002	−0.002	−0.001	−0.002
	(−1.99)	(−2.32)	(−2.21)	(−3.32)	(−2.96)	(−2.24)	(−2.48)	(−2.31)	(−3.42)	(−3.14)
LOGME	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002	−0.001	−0.001	−0.001
	(−2.99)	(−4.40)	(−2.10)	(−2.74)	(−3.29)	(−4.39)	(−5.62)	(−2.72)	(−3.85)	(−4.77)
LOGBM	0.001	0.001	0.000	0.002	0.001	0.001	0.001	0.000	0.002	0.001
	(2.64)	(2.83)	(0.55)	(4.25)	(2.02)	(2.18)	(2.11)	(−0.20)	(4.01)	(1.49)
VNSP	0.005	0.010	0.022	0.022	0.025	0.008	0.006	0.019	0.025	0.023
	(0.98)	(2.51)	(3.82)	(4.97)	(7.03)	(1.68)	(1.29)	(3.05)	(5.67)	(6.84)
IVol	−0.220	−0.053	−0.115	−0.057	−0.154	−0.198	−0.119	−0.164	−0.116	−0.201
	(−5.38)	(−1.96)	(−3.96)	(−1.86)	(−4.93)	(−5.53)	(−4.49)	(−6.2)	(−4.62)	(−7.81)
β	0.002	0.002	0.003	0.002	0.001	0.003	0.003	0.003	0.002	0.002
	(2.30)	(1.53)	(2.20)	(1.42)	(1.29)	(3.49)	(2.51)	(2.87)	(1.95)	(2.53)
Vol						−0.028	−0.027	−0.028	−0.011	−0.016
						(−2.15)	(−2.25)	(−2.03)	(−0.90)	(−1.27)
Turnover	−0.028	−0.023	−0.013	−0.016	−0.025	−0.023	−0.015	−0.008	−0.012	−0.021
	(−1.87)	(−2.09)	(−1.45)	(−1.13)	(−1.65)	(−1.63)	(−1.12)	(−0.78)	(−0.90)	(−1.49)

Notes. This table reports the time series average of the regression coefficients from Fama–MacBeth regressions controlling for the interaction effect of volatility (*Vol*) and CGO. Panel A controls for the interaction effect of *IVol* and CGO. Panel B controls for the interaction effect of *RetVol* and CGO. Variable definitions and sample period are the same as in Table 6. The intercept of the regression is not reported. The *t*-statistics are in parentheses, and they are calculated based on the heteroskedasticity-consistent standard errors of White (1980). *LOGBM* is the log of book to equity, and *LOGME* is the log of market equity.

replace the continuous CGO variable with a high-CGO dummy and a low-CGO dummy, indicating stocks in the top and bottom CGO terciles, respectively. The results remain strong and robust using this more nonparametric characterization of unrealized capital gains, and they are shown in Table IA8 of Online Appendix II.

3.5. Investor Trading Behavior

Our earlier evidence indicates that investors' RDP affects asset prices and especially plays a significant role in the lottery return spreads. In this section, we provide complementary evidence on investors' preference for lottery stocks by directly examining investor trading behavior. More specifically, we investigate investors' preference for lottery-like stocks after gains versus losses among both individual traders and mutual fund managers. Our hypothesis is that investors exhibit stronger preferences for lottery-like assets after losses than after gains.

For individual traders, we use the trading data employed by Barber and Odean (2000). These data come from a large discount brokerage firm, span the time series from January 1991 to December 1996, and consist of 78,000 household accounts, among which we randomly selected 10,000 accounts to conduct our analysis.[25] We follow Ben-David and Hirshleifer (2012) in cleaning the data. Observations are at the investor/stock/day level.

Mutual funds holding data are taken from the Thomson Reuters Mutual Fund and Institutional Holdings databases from the S12 Master Files, which date back to January 1980. We include all U.S. common shares

Table 10. Double Sorts in Subsamples of Top and Bottom Institutional Ownership or Nominal Stock Price

	Top 25% IO			Bottom 25% IO			Top − bottom IO		
Lottery proxy	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1
Maxret	−0.94	0.21	1.15	−2.43	−0.22	2.21	1.49	0.43	−1.06
	(−3.48)	(0.67)	(3.65)	(−8.58)	(−0.68)	(5.15)	(4.23)	(1.15)	(−2.00)
Jackpotp	−0.79	−0.22	0.57	−2.17	−0.19	1.97	1.38	−0.02	−1.40
	(−3.32)	(−0.92)	(1.83)	(−5.56)	(−0.57)	(3.92)	(3.06)	(−0.06)	(−2.37)
Skewexp	−0.79	−0.63	0.15	−1.74	0.42	2.17	0.96	−1.06	−2.01
	(−2.39)	(−2.25)	(0.42)	(−4.92)	(1.37)	(4.84)	(2.14)	(−2.46)	(−3.52)
Deathp	−0.90	−0.52	0.38	−1.78	−0.90	0.88	0.88	0.38	−0.49
	(−2.95)	(−1.81)	(1.07)	(−4.52)	(−2.10)	(1.68)	(1.89)	(0.74)	(−0.79)
Oscorep	−0.42	0.09	0.50	−1.65	0.25	1.90	1.23	−0.16	−1.39
	(−1.54)	(0.45)	(1.68)	(−4.07)	(0.84)	(4.30)	(2.59)	(−0.45)	(−2.68)

	Top 25% price			Bottom 25% price			Top − bottom price		
Lottery proxy	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1	CGO1	CGO5	C5 − C1
Maxret	−0.46	0.62	1.08	−3.00	−1.11	1.88	2.54	1.73	−0.81
	(−2.30)	(3.07)	(4.61)	(−9.5)	(−4.13)	(5.00)	(7.84)	(5.64)	(−1.98)
Jackpotp	−0.42	0.50	0.92	−1.69	−0.37	1.32	1.27	0.87	−0.40
	(−1.99)	(2.60)	(3.74)	(−5.34)	(−1.09)	(3.22)	(3.61)	(2.33)	(−0.87)
Skewexp	−0.86	−0.90	−0.05	−1.00	0.51	1.51	0.14	−1.42	−1.56
	(−2.85)	(−3.77)	(−0.15)	(−2.45)	(1.45)	(2.90)	(0.30)	(−3.79)	(−2.74)
Deathp	−0.61	−0.39	0.21	−1.91	−1.09	0.82	1.31	0.70	−0.61
	(−2.74)	(−1.85)	(0.84)	(−5.84)	(−3.21)	(1.82)	(3.65)	(1.79)	(−1.27)
Oscorep	−0.30	0.09	0.40	−1.72	−0.46	1.26	1.41	0.55	−0.86
	(−1.77)	(0.54)	(1.77)	(−5.93)	(−1.86)	(3.52)	(4.23)	(1.91)	(−2.10)

Notes. This table reports the Fama–French three-factor monthly α values (in percentages) for the lottery spread (difference between top- and bottom-quintile lottery portfolios) among the bottom- and top-tercile CGO portfolios and their differences within top 25% institutional ownership (IO; or nominal stock price) and bottom 25% IO (or nominal stock price) stocks. At the beginning of every month, we first divide stocks into three groups (top 25%, middle 50%, and bottom 25%) by IO (or price), and within each subgroup, stocks are further independently sorted into three groups based on the lagged CGO of Grinblatt and Han (2005) and five groups based on lagged lottery proxies. The portfolio is then held for one month. IO is the percentage of shares held by institutions each month. The CGO of Grinblatt and Han (2005) at week t is computed as one less the ratio of the beginning of the week t reference price to the end of week $t-1$ price. The week t reference price is the average cost basis calculated as $RP_t = k^{-1}\sum_{n=1}^{T}\left(V_{t-n}\prod_{\tau=1}^{n-1}(1-V_{t-n-\tau})\right)P_{t-n}$, where V_t is week t's turnover in the stock, T is the number of weeks in the previous five years, and k is a constant that makes the weights on past prices sum to one. Turnover is calculated as trading volume divided by number of shares outstanding. Monthly CGO is weekly CGO of the last week in each month. We consider five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). In the cases of IO portfolios, the sample period is from January 1980 to October 2014 for *Maxret*, *Oscorep*, *Jackpotp*, and *Deathp* and from January 1988 to October 2014 for *Skewexp*. In the cases of price portfolios, the sample period is from January 1965 to December 2014 for *Maxret* and *Oscorep*, from January 1972 to December 2014 for *Jackpotp* and *Deathp*, and from January 1988 to December 2014 for *Skewexp*. The t-statistics in parentheses are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

with CRSP share codes corresponding to 10 and 11 and apply filters from Frazzini (2006) to exclude erroneous observations.[26] Observations are at the fund/stock/report day level, where funds typically report their holdings at a quarterly frequency. Following this literature, we assume that trading happens on the report date.

We perform probit regressions of a selling indicator on investors' gains and losses (Ret^+ and Ret^-), the lottery feature of a stock, and the interaction between these two as well as other controls. We use the five lottery measures elaborated in the preceding section to proxy for the lottery feature of a stock. For both retail investors and mutual funds trading, we adopt a first-in-first-out assumption in calculating investors' return since purchase. If an investor has made several purchases at various points, we take a weighted average of purchase prices, where the weight equals the percentage of shares bought at that time that are still held by the investor. The terms Ret^+ and Ret^- are the positive and negative parts of the return since purchase, respectively ($Ret^+ = \text{Max}\{Ret, 0\}$ and $Ret^- = \text{Min}\{Ret, 0\}$).

The terms $Proxy \times Ret^+$ and $Proxy \times Ret^-$ are the interaction terms of the lottery feature and gains and losses, where proxy stands for one of these lottery measures. Other control variables include an indicator that equals 1 if Ret is positive and 0 otherwise ($I(Ret > 0)$), an indicator that equals 1 if Ret is 0 and 0 otherwise ($I(Ret = 0)$), return volatility calculated from the daily returns in the past one year ($RetVol$), the logarithm of

Table 11. Propensity to Sell Lottery Stocks, Individual Investors

	I(Selling)				
Proxy	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*
Ret^{+}	0.0007	0.0005	0.0004	0.0005	0.0009
	(4.90)	(4.60)	(3.43)	(5.03)	(8.98)
Ret^{-}	−0.0028	−0.0012	−0.0011	−0.0013	−0.0004
	(−16.97)	(−8.04)	(−5.56)	(−7.99)	(−3.00)
Proxy	0.0088	−0.0400	−0.0010	−0.3593	−0.0020
	(14.14)	(−6.51)	(−11.44)	(−7.12)	(−6.12)
$Ret^{+} \times Proxy$	0.0038	0.0615	0.0013	0.9305	0.0049
	(2.10)	(5.38)	(5.78)	(7.62)	(5.55)
$Ret^{-} \times Proxy$	0.0367	0.0924	0.0015	0.9433	0.0048
	(18.93)	(7.59)	(5.37)	(7.81)	(5.00)
RetVol	0.0431	0.0689	0.0528	0.0583	0.0578
	(20.75)	(26.96)	(24.94)	(25.02)	(24.26)
log(buy price)	0.0005	0.0003	0.0002	0.0003	0.0004
	(10.84)	(7.52)	(5.80)	(8.48)	(9.39)
sqrt(time owned)	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001
	(−38.62)	(−38.37)	(−39.03)	(−39.20)	(−38.96)
$I(Ret > 0)$	0.0010	0.0010	0.0010	0.0010	0.0010
	(17.47)	(18.25)	(17.73)	(17.89)	(17.36)
$I(Ret = 0)$	−0.0001	−0.0000	−0.0001	−0.0000	−0.0001
	(−1.32)	(−0.23)	(−0.63)	(−0.35)	(−0.79)
Observations	25,615,232	23,827,309	25,524,756	25,439,907	22,632,746
Pseudo-R^2	0.0420	0.0419	0.0421	0.0420	0.0420

Notes. This table presents results from probit regressions in which the dependent variable is a dummy equal to one if a stock was sold and zero otherwise. The coefficients reflect the marginal effect on the average stock selling behavior of individual investors. The data set contains the daily holdings of 10,000 retail investors who are randomly selected from 78,000 households with brokerage accounts at a large discount broker from January 1991 to December 1996. Observations are at the investor/stock/day level. The same data set is used in Barber and Odean (2000, 2001, 2002) and, more recently, in Ben-David and Hirshleifer (2012). Ret^{+} (Ret^{-}) is the return since purchase if the return since purchase is positive (negative) and zero otherwise. Return since purchase is defined as the difference between current price and purchase price divided by purchase price (or weighted average price in the case of multiple purchases). The current price is the selling price, price of buying additional shares, or end-of-day price each day. $I_{Ret>0}$ ($I_{Ret=0}$) is a dummy equal to one if the return since purchase is positive (zero) and zero otherwise. *RetVol* is the total volatility of the daily stock returns over the past year. Log(buy price) is the log of purchase price in dollars. Sqrt(time owned) is the square root of the number of days since purchase. We consider five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). Standard errors are clustered at the investor level, and t-statistics are in parentheses.

purchase price (log(*BuyPrice*)), and the square root of the time since purchase (sqrt(*TimeOwned*), where time is measured in units of trading days for retail investors and months for mutual fund managers).

The timing of our regressions is designed as follows. First, all lottery proxies are calculated at a monthly frequency. For retail investors, the selling indicator on one day is regressed on the lottery proxy measured at the end of the previous month. For mutual fund trading, because typical funds report their holdings on a quarterly basis, trading reported at month end t can actually happen from the beginning of month $t-2$ to the end of month t. To have lottery information available at the time of trading, we lag the lottery measure by three months: that is, using a lottery proxy at month end $t-3$ for the selling indicator at month end t.

Tables 11 and 12 present selling regression results for retail investors and mutual funds, respectively. The coefficients for the interaction terms are usually positive and significant, especially for $Proxy \times Ret^{-}$. This finding implies that investors' preference for lottery-like assets over non-lottery-like assets is significantly stronger in the loss region compared than in the gain region. This pattern generally holds for both retail investors and mutual fund managers, and it is robust to our five measures of lottery. This confirms our conjecture about the role of reference points in an investor's preference for lottery-like assets.

4. Conclusion

In this paper, we document that the return spreads between lottery-like assets and non-lottery-like assets

Table 12. Propensity to Sell Lottery Stocks, Mutual Funds

	I(Selling)				
Proxy	*Maxret*	*Jackpotp*	*Skewexp*	*Deathp*	*Oscorep*
Ret^+	0.2730	0.2828	0.2670	0.2693	0.2789
	(42.45)	(44.30)	(45.35)	(52.53)	(58.25)
Ret^-	−0.1913	−0.1872	−0.2046	−0.1443	−0.1579
	(−22.76)	(−21.79)	(−24.48)	(−17.57)	(−22.33)
Proxy	−0.1407	−2.6514	−0.0286	0.9588	−0.2970
	(−3.78)	(−6.17)	(−6.85)	(1.71)	(−9.46)
$Ret^+ \times Proxy$	0.1531	−1.4538	0.0270	11.3785	0.1960
	(1.79)	(−1.83)	(2.72)	(7.38)	(3.40)
$Ret^- \times Proxy$	0.6025	1.7634	0.1197	5.7481	0.4318
	(8.33)	(3.49)	(8.79)	(5.44)	(4.71)
RetVol	−0.4336	−0.0770	−0.4345	−0.9459	−0.8969
	(−3.38)	(−0.49)	(−4.13)	(−8.39)	(−8.17)
log(buy price)	0.0436	0.0352	0.0395	0.0442	0.0371
	(12.06)	(11.62)	(10.89)	(11.26)	(10.98)
sqrt(time owned)	−0.0025	−0.0026	−0.0025	−0.0024	−0.0025
	(−9.20)	(−9.15)	(−8.80)	(−8.68)	(−9.08)
$I(Ret > 0)$	−0.0142	−0.0115	−0.0131	−0.0163	−0.0150
	(−13.88)	(−11.33)	(−12.52)	(−14.80)	(−14.87)
$I(Ret = 0)$	−0.0872	−0.0667	−0.0655	−0.0818	−0.0724
	(−21.80)	(−14.36)	(−13.64)	(−18.77)	(−16.72)
Observations	29,619,224	23,164,195	25,382,915	26,261,635	23,509,029
Pseudo-R^2	0.0132	0.0140	0.0142	0.0130	0.0140

Notes. This table presents results from probit regressions in which the dependent variable is a dummy equal to one if a stock was sold and zero otherwise. The coefficients reflect the marginal effect on the average stock selling of mutual funds. The data set is from the Thomson Reuters S12 Master Files, and the sample period is 1980–2013. Observations are at fund/stock/report day level, where funds typically report their holdings at a quarterly frequency. Following this literature, we assume that trading happens on the report date. Ret^+ (Ret^-) is the return since purchase if the return since purchase is positive (negative) and zero otherwise. Return since purchase is defined as the difference between the current price and the purchase price divided by the purchase price (or weighted average price in the case of multiple purchases). The current price is the selling price, price of buying additional shares, or end-of-day price each day. $I_{Ret>0}$ ($I_{Ret=0}$) is a dummy equal to one if the return since purchase is positive (zero) and zero otherwise. *RetVol* is the total volatility of the daily stock returns over the past year. Log(buy price) is the log of purchase price in dollars. Sqrt(time owned) is the square root of the number of days since purchase. We consider five lottery proxies: *Maxret* is the maximum daily return in the last month, *Jackpotp* is the predicted jackpot probability in the last month from Conrad et al. (2014), *Skewexp* is the expected idiosyncratic skewness in the last month from Boyer et al. (2010), *Deathp* is the predicted failure probability in the last month from Campbell et al. (2008), and *Oscorep* is the predicted bankruptcy probability in the last month from Ohlson (1980). Standard errors are clustered at the fund level, and *t*-statistics are in parentheses.

vary substantially across portfolios with different levels of capital gains or losses. More specifically, the previously documented underperformance of lottery-like assets is significantly stronger among firms with prior capital losses. Among firms where investors face large prior capital gains in these investments, the underperformance of lottery-like assets is either weak or even reversed.

We consider several alternative explanations for this empirical pattern, and we find that reference-dependent demand for lottery-like assets is likely the most plausible one. In particular, the break-even effect and the aversion to loss realization suggest that, after losses, investors often take the chance that can recover their prior losses, and the urge to break even can induce investors with prior losses to take risky gambles that they otherwise would not have taken. Under this preference, assets with high skewness seem especially attractive because they provide a better chance of breaking even. Combined with MA, investors' demand for lottery-like assets is much stronger among stocks where average investors are in losses than among stocks where average investors are in gains, leading to stronger underperformance of lottery-like assets among firms with prior capital losses.

Our empirical findings are robust across five different proxies that are studied in the literature of lottery-related anomalies. It suggests that a common factor may have played a critical role in all of these anomalies and calls for a unified framework to understand them. Although our empirical findings are consistent with RDP based on a static argument, Barberis and Xiong (2009) show that a dynamic setting is important in understanding this issue. It is desirable to develop a formal dynamic model to account for our empirical findings in the future.

Acknowledgments

The authors thank Nick Barberis, Byoung-Hyoun Hwang, Danling Jiang, Lei Jiang, Tse-Chun Lin (discussant), Abhiroop Mukherjee, Jeremy Page (discussant), Alberto Rossi (discussant), Maya Shaton (discussant), Yu Yuan, Guofu Zhou, and seminar participants at the University of Minnesota, Florida State University, PBC School of Finance at Tsinghua University, Peking University, the Dallas Fed, Cheung Kong Graduate School of Business, University of Delaware, University of British Columbia, 2015 Northern Finance Association conference, 2016 China International Conference in Finance, 2016 Financial Intermediation Research Society conference, 2016 European Finance Association conference, and 2016 Drexel–Lehigh–Temple–University of Delaware Research Symposium for helpful comments and discussions. They also thank Terry Odean for providing the brokerage data.

Endnotes

[1] The probability weighting over extreme events has been applied to understand many phenomena in finance, economics, and insurance. For a recent review, see Barberis (2013).

[2] Bali et al. (2011, 2017) also argue that the preference for lottery can account for the puzzle that firms with low volatility and low β tend to earn higher returns.

[3] To clarify, our results do not exclude the existence of overweighting small-probability events. In fact, we find that the negative skewness–return relation is generally significant among stocks around the zero-CGO region, which supports an independent role for probability weighting in the lottery-related anomalies.

[4] In a two-period setting with a cumulative prospect theory preference but without MA, Barberis and Huang (2008) show that the CAPM still holds under assumptions, such as multivariate normal distribution for security payoffs. When there is a violation of these assumptions (e.g., MA or the multivariate normality assumption for security payoffs), the CAPM typically fails.

[5] Several studies also apply the reference-dependent feature in decision making to understand various other empirical findings in financial data. See Baker et al. (2012) for information on merger/acquisitions, George and Hwang (2004) and Li and Yu (2012) for information on the predictive power of 52-week high prices, and Dougal et al. (2015) for information on the credit spread.

[6] Our approach is reminiscent of the studies on habit formation. Campbell and Cochrane (1999) show that external habit formation can help account for the equity premium puzzle. In the following studies, Wachter (2006) and Verdelhan (2010) find that the same mechanism can account for the bond return predictability and the forward premium puzzle, respectively. These subsequent studies thus further validate the role of habit formation on asset price dynamics.

[7] For details, see equations 9 and 11 in Grinblatt and Han (2005).

[8] See equations 1 and 2 in Frazzini (2006) for details.

[9] Although our prior is that the lottery preference should be stronger among retail investors, as documented in Kumar (2009), this preference does not have to be confined to retail investors. A growing literature has shown that mutual fund managers exhibit many behavioral biases just like retail investors do. For instance, they exhibit the disposition effect (Frazzini 2006, An and Argyle 2017) and the rank effect (Hartzmark 2015), and they have rolling mental accounts (Frydman et al. 2018). Even professional traders have exhibited loss aversion (Coval and Shumway 2005). DeVault et al. (2019) argue that many institutional investors could be sentiment traders. Agarwal et al. (2018) show that mutual funds that are smaller and younger with poorer recent performance and more retail clientele tend to hold more lottery stocks, which could be associated with incentives to attract capital.

[10] *Deathp* and *Oscorep* are initially motivated to study firms' distress risk. Serving as a proxy for lottery feature is one interpretation among many that have been put forth to explain the negative relation between these measures and future returns.

[11] Related to this finding, Stambaugh et al. (2012) find that many anomalies are driven by the abnormally low returns from their short legs, especially after high-sentiment periods. They argue that this evidence is consistent with the notion that overpricing is more prevalent than underpricing because of short-sale impediments.

[12] In a recent study, Jiang et al. (2016) use different measures of skewness, and they also find that the negative return spread between firms with low and high skewness is more pronounced among firms with low CGO than among firms with high CGO.

[13] We thank the referee for encouraging us to investigate this positive α among high-CGO firms.

[14] Recently, Belo et al. (2014) also emphasized the importance of reporting both equal- and value-weighted portfolio returns.

[15] Another feature of prospect theory is that investors tend to overweight small-probability events. The asset pricing implications of probability weighting have been studied recently by Barberis and Huang (2008), Bali et al. (2011), Green and Hwang (2012), and Barberis et al. (2016), among others.

[16] See, for example, Shefrin and Statman (1985), Benartzi and Thaler (1995), Odean (1998), Barberis et al. (2001), Grinblatt and Han (2005), Frazzini (2006), and Barberis and Xiong (2012), among others.

[17] Once again, we acknowledge that our static argument above may not be valid in a dynamic setting, as shown by Barberis and Xiong (2009). Thus, before fully embracing our argument, one should develop a fully dynamic model, which is beyond the scope of our study. See Li and Yang (2013) for such a related dynamic model.

[18] We thank Terry Odean for the brokerage data.

[19] However, by exploring crosscountry variation in creditor protection, Gao et al. (2017) argue that shareholder expropriation is unlikely to account for the distress anomaly.

[20] For recent evidence on how risk attitude is affected by realized versus unrealized profits, see Imas (2016).

[21] In Table IA7 in Online Appendix II, we show that our results remain similar when we replace past returns with other proxies for news, including the most recent available standardized unexpected earnings, and cumulative abnormal returns around the most recent earnings announcement.

[22] For example, Avramov et al. (2013) show that many anomalies are only significant among distressed firms, suggesting that distressed firms are more difficult to arbitrage.

[23] One could use the nominal price level as another proxy for the lottery feature, as in Kumar (2009). Indeed, in untabulated analysis, we find that our results hold well when the nominal price is used as a lottery proxy.

[24] Recently, in a related paper, Lin and Liu (2016) find that the lottery-related anomalies are more pronounced among firms with stronger individual demand.

[25] Because of computational limitations, randomly selecting a sample of 10,000 is a general convention among studies using this data set. See, for example, Odean (1998) and Ben-David and Hirshleifer (2012).

[26] Observations are excluded if (1) the number of shares in a fund's portfolio is greater than the total number of shares outstanding in that stock, (2) the value of the fund holding of one stock is greater than the total asset value of the fund, (3) the stock has zero shares outstanding, and (4) the value of a fund reported by Thomson Reuters is different from the implied CRSP value by more than 100%.

References

Agarwal V, Jiang L, Wen Q (2018) Why do mutual funds hold lottery stocks? Working paper, Georgia State University, Atlanta.

Amihud Y (2002) Illiquidity and stock returns: Cross-section and time-series effects. *J. Financial Markets* 5(1):31–56.

An L (2016) Asset pricing when traders sell extreme winners and losers. *Rev. Financial Stud.* 29(3):823–861.

An L, Argyle B (2017) Overselling winners and losers: Mutual fund trading behavior and price effects. Working paper, Tsinghua University, Beijing.

Anderson A-M, DylEA (2005) Market structure and trading volume. *J. Financial Res.* 28(1):115–131.

Asparouhova E, Bessembinder H, Kalcheva I (2013) Noisy prices and inference regarding returns. *J. Finance* 68(2):665–714.

Avramov D, Chordia T, Jostova G, Alexander P (2013) Anomalies and financial distress. *J. Financial Econom.* 108(1):139–159.

Baker M, Wurgler J (2006) Investor sentiment and the cross-section of stock returns. *J. Finance* 61(4):1645–1680.

Baker M, Pan X, Wurgler J (2012) The effect of reference point prices on mergers and acquisitions. *J. Financial Econom.* 106(1):49–71.

Bali TG, Cakici N (2008) Idiosyncratic volatility and the cross section of expected returns. *J. Financial Quant. Anal.* 43(1):29–58.

Bali TG, Murray S (2013) Does risk-neutral skewness predict the cross section of equity option portfolio returns? *J. Financial Quant. Anal.* 48(4):1145–1171.

Bali TG, Cakici N, Whitelaw RF (2011) Maxing out: Stocks as lotteries and the cross-section of expected returns. *J. Financial Econom.* 99(2):427–446.

Bali TG, Brown SJ, Murray S, Tang Yi (2017) A lottery demand-based explanation of the beta anomaly. *J. Financial Quant. Anal.* 52(6): 2369–2397.

Bali TG, Cakici N, Yan X, Zhang Z (2005) Does idiosyncratic risk really matter? *J. Finance* 60(2):905–929.

Barber BM, Odean T (2000) Trading is hazardous to your wealth: The common stock investment performance of individual investors. *J. Finance* 55(2):773–806.

Barber BM, Odean T (2001) Boys will be boys: Gender, overconfidence, and common stock investment. *Quart. J. Econom.* 116(1):261–292.

Barber BM, Odean T (2002) Online investors: Do the slow die first? *Rev. Financial Stud.* 15(2):455–488.

Barberis N (2013) The psychology of tail events: Progress and challenges. *Amer. Econom. Rev.* 103(3):611–616.

Barberis N, Huang M (2001) Mental accounting, loss aversion, and individual stock returns. *J. Finance* 56(4):1247–1292.

Barberis N, Huang M (2008) Stocks as lotteries: The implications of probability weighting for security prices. *Amer. Econom. Rev.* 98(5):2066–2100.

Barberis N, Xiong W (2009) What drives the disposition effect? An analysis of a long-standing preference-based explanation. *J. Finance* 64(2):751–784.

Barberis N, Xiong W (2012) Realization utility. *J. Financial Econom.* 104(2):251–271.

Barberis N, Huang M, Santos T (2001) Prospect theory and asset prices. *Quart. J. Econom.* 116(1):1–53.

Barberis N, Mukherjee A, Wang B (2016) Prospect theory and stock returns: An empirical test. *Rev. Financial Stud.* 29(11):3068–3107.

Belo F, Lin X, Santiago B (2014) Labor hiring, investment, and stock return predictability in the cross section. *J. Political Econom.* 122(1): 129–177.

Benartzi S, Thaler RH (1995) Myopic loss aversion and the equity premium puzzle. *Quart. J. Econom.* 110(1):73–92.

Ben-David I, Hirshleifer D (2012) Are investors really reluctant to realize their losses? Trading responses to past returns and the disposition effect. *Rev. Financial Stud.* 25(8):2485–2532.

Boyer B, Todd M, Keith V (2010) Expected idiosyncratic skewness. *Rev. Financial Stud.* 23(1):169–202.

Campbell JY, Cochrane JH (1999) By force of habit: A consumption-based explanation of aggregate stock market behavior. *J. Political Econom.* 107(2):205–251.

Campbell JY, Hilscher J, Szilagyi J (2008) In search of distress risk. *J. Finance* 63(6):2899–2939.

Conrad J, Dittmar RF, Ghysels E (2013) Ex ante skewness and expected stock returns. *J. Finance* 68(1):85–124.

Conrad J, Kapadia N, Xing Y (2014) Death and jackpot: Why do individual investors hold overpriced stocks? *J. Financial Econom.* 113(3):455–475.

Coval JD, Shumway T (2005) Do behavioral biases affect prices? *J. Finance* 60(1):1–34.

Daniel K, Hirshleifer D, Subrahmanyam A (1998) Investor psychology and security market under- and overreactions. *J. Finance* 53(6):1839–1885.

Daniel K, Hirshleifer D, Subrahmanyam A (2001) Overconfidence, arbitrage, and equilibrium asset pricing. *J. Finance* 56(3):921–965.

DeVault L, Sias R, Starks L (2019) Sentiment metrics and investor demand. *J. Finance*. Forthcoming.

Dhar R, Zhu N (2006) Up close and personal: Investor sophistication and the disposition effect. *Management Sci.* 52(5):726–740.

Dougal C, Joseph E, Parsons CA, Van Wesep ED (2015) Anchoring on credit spreads. *J. Finance* 70(3):1039–1080.

Falkenstein EG (1996) Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *J. Finance* 51(1):111–135.

Fama EF, French KR (2008) Dissecting anomalies. *J. Finance* 63(4): 1653–1678.

Fama EF, MacBeth JD (1973) Risk, return, and equilibrium: Empirical tests. *J. Political Econom.* 81(3):607–636.

Frazzini A (2006) The disposition effect and underreaction to news. *J. Finance* 61(4):2017–2046.

Frederick S, Loewenstein G (1999) Hedonic adaptation. Kahneman D, Diener E, Schwartz N, eds. *Scientific Perspectives on Enjoyment, Suffering, and Well-Being* (Russell Sage Foundation, New York), 302–329.

Frydman C, Hartzmark S, Solomon D (2018) Rolling mental accounts. *Rev. Financial Stud.* 31(1):362–397.

Gao P, Parsons CA, Shen J (2017) The global relation between financial distress and equity returns. *Rev. Financial Stud.* 31(1):239–277.

Garlappi L, Yan H (2011) Financial distress and the cross-section of equity returns. *J. Finance* 66(3):789–822.

Garlappi L, Shu T, Yan H (2008) Default risk, shareholder advantage, and stock returns. *Rev. Financial Stud.* 21(6):2743–2778.

George TJ, Hwang C-Y (2004) The 52-week high and momentum investing. *J. Finance* 59(5):2145–2176.

Gompers PA, Metrick A (2001) Institutional investors and equity prices. *Quart. J. Econom.* 116(1):229–259.

Green TC, Hwang B-H (2012) Initial public offerings as lotteries: Skewness preference and first-day returns. *Management Sci.* 58(2): 432–444.

Grinblatt M, Han B (2005) Prospect theory, mental accounting, and momentum. *J. Financial Econom.* 78(2):311–339.

Grinblatt M, Keloharju M (2001) What makes investors trade? *J. Finance* 56(2):589–616.

Hartzmark S (2015) The worst, the best, ignoring all the rest: The rank effect and trading behavior. *Rev. Financial Stud.* 28(4): 1024–9659.

Imas A (2016) The realization effect: Risk-taking after realized versus paper losses. *Amer. Econom. Rev.* 106(8):2086–2109.

Ingersoll JE Jr, Jin LJ (2013) Realization utility with reference-dependent preferences. *Rev. Financial Stud.* 26(3):723–767.

Jiang L, Wu K, Zhou G, Zhu Y (2016) Stock return asymmetry: Beyond skewness. Working paper, Tsinghua University, Beijing.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.

Kumar A (2009) Who gambles in the stock market? *J. Finance* 64(4): 1889–1933.

Kumar A, Page JK, Spalt OG (2011) Religious beliefs, gambling attitudes, and financial market outcomes. *J. Financial Econom.* 102(3):671–708.

Li J, Yu J (2012) Investor attention, psychological anchors, and stock return predictability. *J. Financial Econom.* 104(2):401–419.

Li Y, Yang L (2013) Prospect theory, the disposition effect, and asset prices. *J. Financial Econom.* 107(3):715–739.

Lin T-C, Liu X (2016) Skewness, individual investor preference, and the cross-section of stock returns. Working paper, Hong Kong University, Hong Kong.

Locke PR, Mann SC (2000) Do professional traders exhibit loss realization aversion? Working paper, George Washington University, Washington, DC.

Nagel S (2005) Short sales, institutional investors and the cross-section of stock returns. *J. Financial Econom.* 78(2):277–309.

Newey WK, West KD (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3):703–708.

Odean T (1998) Are investors reluctant to realize their losses? *J. Finance* 53(5):1775–1798.

Ohlson JA (1980) Financial ratios and the probabilistic prediction of bankruptcy. *J. Accounting Res.* 18(1):109–131.

Rayo L, Becker GS (2007) Evolutionary efficiency and happiness. *J. Political Econom.* 115(2):302–337.

Shapira Z, Venezia I (2001) Patterns of behavior of professionally managed and independent investors. *J. Banking Finance* 25(8): 1573–1587.

Shefrin H, Statman M (1985) The disposition to sell winners too early and ride losers too long: Theory and evidence. *J. Finance* 40(3): 777–790.

Stambaugh RF, Yu J, Yuan Y (2012) The short of it: Investor sentiment and anomalies. *J. Financial Econom.* 104(2):288–302.

Thaler R (1980) Toward a positive theory of consumer choice. *J. Econom. Behav. Organ.* 1(1):39–60.

Thaler R (1985) Mental accounting and consumer choice. *Marketing Sci.* 4(3):199–214.

Thaler RH, Johnson EJ (1990) Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Sci.* 36(6):643–660.

Verdelhan A (2010) A habit-based explanation of the exchange rate risk premium. *J. Finance* 65(1):123–146.

Wachter JA (2006) A consumption-based model of the term structure of interest rates. *J. Financial Econom.* 79(2):365–399.

Wang H, Yan J, Yu J (2017) Reference-dependent preferences and the risk-return trade-off. *J. Financial Econom.* 123(2): 395–414.

Wermers R (2003) Is money really "smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working paper, University of Maryland, College Park.

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48(4):817–838.

Xing Y, Zhang X, Zhao R (2010) What does the individual option volatility smirk tell us about future equity returns? *J. Financial Quant. Anal.* 45(3):641–662.

Zhang XF (2006) Information uncertainty and stock returns. *J. Finance* 61(1):105–137.

Zipf GK (1949) *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Reading, MA).