Lottery-Related Anomalies:
The Role of Reference-Dependent Preferences[*]

Li An
Tsinghua University

Huijun Wang
University of Delaware

Jian Wang
Federal Reserve Bank of Dallas

Jianfeng Yu
University of Minnesota

December 2015

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, while 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 united framework to understand the lottery-related anomalies in the literature.

JEL codes: G02, G12, G14

1 Introduction

Numerous studies have found that lottery-like stocks tend to significantly underperform non-lottery-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. 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, though 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 employing Grinblatt and Han's (2005) method 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 with top quintile of maximum daily returns in the previous month) are 138 basis points (bps) lower per month than non-lottery-like firms (those with

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 non-lottery-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 Fama-MacBeth regressions.

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 though 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 it calls for a unified framework to understand these anomalies. Therefore, we go on to examine several possible explanations for our empirical findings. First, we investigate the role of reference-dependent preferences (RDP) 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 following prior losses relative to a reference point (the break-even effect). Lottery-like assets are particularly attractive in these cases since they provide a better chance to recover prior losses. On the other hand, when investors face prior gains, their demand for lottery-like assets is not as strong since they are not risk seeking or in need of breaking even. Instead, due to 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 to non-lottery-like stocks among the stocks where investors face prior losses, leading to lower future returns than non-lottery-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 since investors with capital gains usually

2

dislike the high volatility of lottery-like stocks. Thus, RDP together with MA can potentially account for the empirical findings documented in this paper. However, we acknowledge that the static argument here might not be valid in a dynamic setting as shown in Barberis and Xiong (2009). It would be helpful to develop a formal model in a dynamic setting to account for our empirical findings, which is beyond the scope of this paper and we leave it for future research.

The 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 information travels slowly across investors, stocks with low CGO tend to be overvalued on average. Moreover, information is very likely to travel more slowly among firms with higher failure probability since information uncertainty is usually higher and arbitrage forces are more limited for these firms. 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). On the other hand, firms with capital gains (high CGO) have probably experienced good news and therefore have been underpriced due to 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.

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). Due to the disposition effect (i.e., investors' tendency to sell securities whose prices have increased since purchase rather than those whose prices have dropped), firms with higher CGO experience greater selling pressure and thus are underpriced. Since 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, since 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, while 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-MacBeth regressions to control for: 1) the interaction terms of our lottery proxies and a proxy for past news; 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 following 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 employing different subsamples, such as excluding NASDAQ 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 since more individuals are investing in these stocks. A similarly stronger effect is observed following 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, Hilscher, and Szilagyi (2008) show that firms with a high probability of default have abnormally low average future returns. Conrad, Kapaida, and Xing (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, Mitton, and Vorkink (2010) find that expected idiosyncratic skewness and future returns are negatively correlated. Bali, Cakici, and Whitelaw (2011) show that maximum daily returns in the past month are negatively associated with future returns.[1] All of these empirical studies suggest that positively skewed stocks can be overpriced and earn lower future returns. In addition, several studies have employed option data to study the relation between various skewness measures and future returns of options. For instance, see Xing, Zhang, and Zhao (2010), Bali and Murray (2013), and Conrad, Dittmar, and Ghysels (2013).

---

[1]Bali, Cakici, and Whitelaw (2011) and Bali, Brown, Murray and Tang (2014) also argue that the preference for lottery can account for the puzzle that firms with low volatility and low beta tend to earn higher returns.

We differ from the above studies by showing that the negative skewness-return relation are 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.[2] In particular, we find supportive evidence for RDP being an important source for lottery-related anomalies besides 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, Huang, and Santos (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.[3] 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 that are studied in these theoretical papers.[4]

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, Yan, and Yu (2014) show that RDP may have also played an important role in the lack of positive risk-return trade-off in the data. We show that the effect of CGO on lottery-related anomalies is distinct from the effect of CGO on the risk return trade-off and is not primarily driven by investor's RDP for volatility risk although lottery-like assets tend to have higher volatility. In particular, we employ residual skewness measures which are orthogonal to volatility, and we find a similar effect of CGO on the residual skewness-return

---

[2]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 zero-CGO region, which supports an independent role for probability weighting in the lottery-related anomalies.

[3]In a two-period setting with a cumulative prospect theory preference but without mental accounting, 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., mental accounting or the multivariate normality assumption for security payoffs), the CAPM typically fails.

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

relation.

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 is dependent 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 the Fama-MacBeth regressions results are reported.

Our data are obtained from several sources. Stock data are from monthly and daily CRSP database, accounting data are from Compustat Annually and Quarterly database, and mutual fund holdings data are obtained from the Thomson Financial CDA/Spectrum Mutual Funds database. To construct stock-level variables, we start with the data of all US common stocks traded in NYSE, AMEX, and NASDAQ from 1962 to 2014. Next, we filter our dataset by requiring all observations to have nonnegative book equity, prices to be equal to or greater than \$5, and to have at least 10 non-missing 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 Appendix II.

2.1.1 Capital Gains Overhang

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.[5] 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. Since 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 non-missing 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).[6] 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 employs 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, due to 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.

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 and more details on the construction of these measures are provided in Appendix II.

Maxret: Bali, Cakici, and Whitelaw (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 due to 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

---

[5]For details, see Equation (9) on page 319 and Equation (11) on page 320 in Grinblatt and Han (2005).
[6]See Equation (1) on page 2022 and Equation (2) on page 2023 of Frazzini (2006) for details.

of lottery feature.

Jackpotp: Conrad, Kapadia, and Xing (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 which are positively skewed. Thus, we use the predicted probability of jackpot (log returns greater than 100% over the next year) which is estimated from their baseline model (Panel A of Table 3 on page 461) as our second measure. The out-of-sample predicted jackpot probabilities start from January 1972 in our paper.

Skewexp: Boyer, Mitton, and Vorkink (2010) estimate a cross-sectional model of expected idiosyncratic skewness and find that it negatively predicts future returns. We use the expected idiosyncratic skewness estimated from their model (model 6 of Table 2 on page 179) as our third measure. Due to the limited availability of NASDAQ turnover data in earlier years, this measure starts in January 1988.

Deathp: Campbell, Hilscher, and Szilagyi (2008) find that stocks with a high predicted failure probability earn abysmally low subsequent returns. Since distressed stocks tend to have positive skewness, they conjecture that investors have strong preference for positive skewness which drives up the prices of distress stocks and leads to lower future returns. We construct this proxy as our fourth measure of lottery feature, using their logit model (Table IV, 12 month lag on page 2913). The sample period of *Deathp* starts in January 1972 due to 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 higher bankruptcy probability earn lower subsequent returns. Following his approach, we calculate firms' predicted bankruptcy probability based on the O-score (Model 1 of Table 4 on page 121) and use this proxy as our fifth measure of 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. 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) or one of the five lottery proxies (Panel B). In each quintile, the portfolio return (Ret^e) is calculated as value-weighted returns of individual stocks in month $t + 1$. α_{FF3} is the value-weighted average of the intercepts of the Fama-French three-factor regression. 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 the ex-post skewness is measured by the 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 Newey-West (1987) standard errors with a lag of 36 for firm characteristics.

Panel A of Table 1 reports summary statistics for portfolios sorted on CGO, using both Grinblatt and Han's (2005) measure and Frazzini's (2006) measure. Consistent with the previous literature, high-CGO firms tend to have larger firm size, higher book-to-market ratios, 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 top and bottom quintiles is 18 basis points per month. In addition, the spread between the Fama-French three-factor alphas for the high- and low-CGO portfolios is 37 bps for Grinblatt and Han's (2005) measure and 39 bps for Frazzini's (2006) measure. 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 portfolios are equally weighted.

Panel B of Table 1 presents monthly excess returns and the Fama-French three-factor alphas for portfolios sorted on the lottery proxies. Consistent with previous studies on each of these anomalies, lottery-like portfolios (row P5) underperform non-lottery-like portfolios (row P1), and the return difference is significant, especially in terms of the Fama-French three-factor alphas. For instance, the Fama-French three-factor alpha spread between 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 also reports ex-post skewness for each portfolio, which is measured by the time series mean of cross-sectional average stock-level skewness calculated from daily stock returns in the next 12 months. As expected, we usually find ex-post skewness increases monotonically from non-lottery-like (P1) portfolios to lottery-like (P5) portfolios for all five lottery proxies. For instance, if the lottery feature is measured by the predicted jackpot probability, the ex-post skewness increases from 0.17 for P1 to 0.60 for P5. The difference between 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.

2.3 Double Sorts

As shown in the previous subsection, our five lottery measures *unconditionally* predict future returns in a way consistent with previous studies in the literature. We now examine to what extent these predictive patterns depend on stocks' previous capital gains/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 Grinblatt and Han's (2005) CGO and the five proxies for lottery-like feature. Panel A reports excess returns for these portfolios, while Panel B presents the Fama-French three-factor alphas.[7] 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. Following previous losses (CGO1), high-Maxret stocks *underperform* low-Maxret stocks by 1.38% per month in excess returns, with the t-stat equal to -5.35. In contrast, following 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-stat 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 (in Table 1) with the t-stat equal to -1.07. Columns C5-C1 report 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-stat of 7.50.

---

[7]The excess return is measured by stock returns minus one-month Treasury bill rate.

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 holds for the Fama-French three-factor alphas, as shown in Panel B.[8] More interestingly, Panel B shows that among low-CGO firms, a large bulk of the return spreads between low- and high-skewness firms is due to the negative alpha of the lottery-like assets. Taking Maxret as an example, the long-leg has an alpha of 0.52% per month, whereas the short-leg has an alpha of -1.24% per month.[9] This is consistent with the notion that facing prior losses, the demand for lottery-like assets increases. Due to 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 out of these five return spreads are positive and three of them are significant. The patterns are similar for the Fama-French three-factor alphas with three out of five spreads being at least marginally significant and the other two negative spreads being insignificant.

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 CGO close to zero. This finding suggests that besides the effect of investors' stronger demand for lottery-like assets following 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 Grinblatt and Han's (2005) CGO is based on price-volume approximation and could be affected by high-frequency trading volume, we employ Frazzini's (2006) CGO, which is based on actual holdings of mutual funds. We repeat the double sorting exercise after replacing Grinblatt and Han's (2005) CGO with Frazzini's (2006) CGO. The results are reported in Table 3, and are very similar to those in Table 2. For example, Panel A shows that the differences between excess return spreads among high-CGO firms

---

[8]Untabulated results show that CAPM alphas and Carhart four-factor alphas have similar patterns.

[9]Related to this finding, Stambaugh, Yu, and Yuan (2012) find that many anomalies are driven by the abnormally low returns from their short-legs, especially following high sentiment periods. They argue that this evidence is consistent with the notion that overpricing is more prevalent than underpricing due to short-sale impediments.

and those among low-CGO firms (C5-C1) are respectively 1.88%, 1.26%, 0.56%, 1.10%, and 0.69% per month with corresponding t-statistics of 5.99, 4.09, 1.55, 3.10, and 2.38 for the five lottery-feature proxies. The sample period in Table 3 is shorter due to 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.

In Panel B of Table 3, the lottery-like and non-lottery-like spreads of alphas 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 alpha spread between low- and high-skewness firms is only -26 bps (v.s. an average spread of -161 bps among low-CGO firms). Thus, the evidence based on Frazzini's CGO confirms that there is virtually no return spreads between lottery-like assets and non-lottery-like assets among firms with large capital gains (high CGO).

There is one caveat of using the raw CGO measure: since 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 other effects rather than the capital gains/losses that investors face. To address this concern, we sort stocks based on the residual CGO (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 alpha spreads between lottery and non-lottery portfolios (P5-P1) for low- and high-RCGO groups in the two right panels. To facilitate comparison, we also include lottery spreads based on raw CGO in the two left panels, which serve as a summary of the results presented in Tables 2 and 3. For each of the five lottery proxies, Panle CGO^{GH} reports the lottery spreads (P5-P1 based on the lottery proxy) among firms with low CGO (CGO1), the lottery spread among firms with high CGO (CGO5), and the difference between these two spreads (C5-C1). In this panel, CGO is based on Grinblatt and Han's (2005) measure. Panel CGO^{FR} presents similar results for CGO calculated from Frazzini's (2006) procedure. The two right panels report the results for RCGO under these two measures of CGO. Using the residual rather than the raw CGO delivers similar results that support our hypothesis as well. Taking RCGO under Grinblatt and Han's (2005) procedure for instance, the difference between the lottery 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 of 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.

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. It helps to avoid our results being dominated by the behavior of very small firms, which as warned by Fama and French (2008). But on the other hand, 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- and lagged-gross-return-weighted portfolio alphas.[10] 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, Bessembinder, and Kalcheva, 2013).

The results in Table 5 confirm a significant role of 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 of 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 return in our benchmark results, suggesting that our findings are not mainly driven by extremely large or small firms.

In the last panel (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 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

---

[10]Recently, Belo, Lin and Bazdresch (2014) also emphasize the importance of reporting both equal- and value-weighted portfolio returns.

panels of Table 5, the results are based on Grinblatt and Han's (2005) CGO measure. The results based on Frazzini's (2006) measure are quantitatively similar and are not reported to save space.

2.4   Fama-MacBeth Regressions

The double-sorting approach in the previous 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 of the Fama-MacBeth regressions below, we control for a list of traditional return predictors, such as firm size, book-to-market, past returns, stock return volatility, and share turnover. Following Conard, Kapadia, and Xing (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 Grinblatt and Han's (2005) finding.   Grinblatt and Han (2005) attribute this finding to investors' tendency of selling stocks with capital gains (high CGO). The 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) to (4) in the next section. Under each lottery proxy, the regression 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, past returns, stock return volatility, and shares turnover. It is noteworthy that the coefficient of lottery proxy itself typically appears 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 return-skewness

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/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 to incorporate 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 due to the disposition effect, rather than investors' status of gains/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.

3.1 The Role of RDP

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

However, RDP has recently attracted massive attentions 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. 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.[11] Finally, the mental

---

[11]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, Cakici, and Whitelaw (2011), and Barberis, Mukherjee, and Wang (2014), among others.

accounting of Thaler (1980, 1985) provides a theoretical foundation for decision makers setting a separate reference point for each asset they own by ignoring possible interactions among these assets.

Building on the RDP model by Kahneman and Tversky (1979) and mental accounting, a large number of recent studies have demonstrated that RDP can better capture human behaviors in many decision making processes and it can account for many asset pricing phenomena that contradict the prediction of standard models[12]. Moreover, psychological and evolutionary foundations for RDP are also documented in Frederick and Loewenstein (1999) and Rayo and Becker (2007).

Among studies suggesting 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) on professional traders in Israel, Wermers (2003) and Frazzini (2006) on mutual fund managers, and Coval and Shumway (2005) on professional market makers at the Chicago Board of Trade.

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 following losses, investors often have strong urge to make up their losses since by breaking even, investors can avoid proving 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 appear especially attractive since they provide a better chance to break even.

In contrast, among stocks with prior capital gains, there are two countervailing forces. On one hand, investors might still prefer lottery-like stocks, probably due to the overweighting of small-probability event in the standard probability weighting scheme of the prospect theory, though the demand for lottery-like assets becomes weaker as the effects from break-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

---

[12] See, e.g., Shefrin and Statman (1985), Benartzi and Thaler (1995), Odean (1998), Barberis, Huang, and Santos (2001), Grinblatt and Han (2005), Frazzini (2006), and Barberis and Xiong (2012), among others.

stock-level idiosyncratic volatility due to mental accounting. Thus, the lottery-like stock can be undervalued and exhibit high future returns. Overall, it is not clear which force dominates in the data. But we can at least conclude from the above discussions that investors' demand for lottery-like stocks should be stronger in the loss region than that in the gain region.

In sum, a natural implication from RDP and mental accounting 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 been seeking break-even opportunities.[13]

Since CGO measures the unrealized gains and losses from investment, the lottery-related anomalies should crucially depend on individual stock' CGO: a strong negative correlation between expected (abnormal) returns and skewness should exist among firms with a low (negative) CGO, while a weak (insignificant or even reversed) correlation between expected abnormal returns and skewness may exist among firms with a high (positive) CGO. Furthermore, the return spreads (between high- and low-skewness stocks) should be significantly more negative among firms with capital losses than those among firms with capital gains.

This is exactly the pattern presented in Section 2. In fact, 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-assets is significantly stronger in the loss region than in the gain region.[14] Using probit regressions, we estimate the propensity to sell lottery-like stocks for individual investors. The results are only reported in the appendix to save space. The coefficients for the interaction terms between unrealized returns and skewness proxies are significant in Tables A1 of the appendix, 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 and the results are reported in Table A2 in the appendix. These results confirm our conjecture about the role of RDP in the lottery anomalies.

Lastly, we discuss the relation between RDP and some other popular explanations in the literature for the documented lottery-related anomalies. The overweighting of small-

---

[13]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 current study.

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

probability events in the 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 the 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 since the default option is a hedge (e.g., Garlappi, Shu, and Yan, 2008 and Garlappi and Yan, 2011).[15]

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 homogenous 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, Cakici, and Whitelaw, 2011, and Conrad, Kapadia, and Xing, 2014). Our empirical findings suggest that RDP may have played a crucial role in account 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.

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

---

[15]However, by exploring cross-country variation in creditor protection, Gao, Parsons and Shen (2014) argue that shareholder expropriation is unlikely to account for the distress anomaly.

be stronger among firms with higher information uncertainty. Thus, among the firms with recent bad news, higher information uncertainty is likely to forecast lower future returns due to the current underreaction to the past bad news.

Our proxies for the lottery-like feature could be related to information uncertainty, especially for Campbell, Hilscher, and Szilagyi's (2008) failure probability and Ohlson's (1980) bankruptcy probability since these firms might indeed be hard to evaluate. Since 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 due to news underreaction. This overpricing effect is more pronounced for lottery-like stocks due to higher information uncertainty, implying a negative relation between the lottery proxies and future returns among firms with low CGO. The above argument is consistent to the return-skewness-CGO pattern observed in Table 2 and Table 3, and 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 employed as a proxy for news. 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 of 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 is against the underreaction-to-news effect being an explanation to 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-statistic for the interaction term is 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

Besides being a proxy for aggregate capital gains/losses, CGO may be also 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 due to the disposition effect (investors being more likely to sell a security upon a gain rather than a loss), which in turn leads to lower current prices and higher future returns. In general, the final mispricing effect 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, leading to a positive coefficient for the interaction term between CGO and skewness proxies in Fama-MacBeth regressions as we have documented. Indeed, one may expect that firms close to default should impose higher arbitrage risk for arbitrageurs.[16] Note that this explanation does not rely on investors having an especially strong preference for lottery-like assets when facing prior losses. It just requires that the skewness proxies are related to limits to arbitrage and CGO itself is associated with mispricng.

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 (forthcoming). 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 (forthcoming) 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 3 out of the 5 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 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 $Proxy \times CGO$

---

[16]For example, Avramov, Chordia, Jostova, and Philipov (2013) show that many anomalies are only significant among distressed firms, suggesting that distressed firms are more difficult to arbitrage.

only change marginally in magnitude and they remain statistically significant for all lottery proxies.

In sum, both the underreaction-to-news effect and the mispricing story cannot account for the return-skewness pattern that we have documented in Table 2. Coupled with the investors' trading behaviors documented in Tables A1 and A2, 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 equally-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, it 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 in column (4) in Table 6.

The second concern is that our empirical findings could be driven by NASDAQ or illiquid stocks. Previous studies (e.g., Bali, Cakici, Yan, and Zhang, 2005) show that some asset pricing phenomena disappear once 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's (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 with the following alternative specifications: 1) We employ 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 NYSE and AMEX; 3) We exclude the most illiquid stocks - those that fall into the top illiquid decile in each month (using Amihud's (2002) illiquidity measure). 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 of the t-statistics remain statistically significant at the 5% level. In addition, Table A3 in the Appendix 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, NASDAQ stocks, or disproportionately affected by small firms, since both the statistical significance and the economic magnitude remain largely the same after controlling for these factors.

Next, we confirm that our results are not mainly driven by investors' reference-dependent preference for return volatility. Since high-skewness stocks are typically also more volatile, it is possible that the underperformance of lottery-like assets among firms with negative CGO is due to investors' preference for volatility (rather than skewness) after losses. For example, the prospect theory posits that investors are risk-seeking after losses, and thus they might prefer stocks with high volatility after losses. Indeed, Wang, Yan, and Yu (2015) find a significant and negative risk-return relation among low-CGO stocks where investors face losses. To ensure that our results are not primarily driven by investors' preference for volatility after losses, we reexamine the patterns on lottery portfolios by purging out the confounding effect from volatility. 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 use the residual lottery proxies to repeat our double sorting exercises. The results reported in Table 8 indicate that the pattern of the lottery spread holds reasonably well when the residual lottery measures are used. In particular, the differences of residual lottery spreads among high- and low-CGO firms are at least marginally significant for all of the lottery proxies except for Oscorep.

Moreover, we also perform Fama-MacBeth regressions to control for the interaction effect between CGO and volatility and other variables and the results are presented in Table 9. After adding the interaction term between CGO and the return volatility into the regressions, the coefficients of $Proxy \times CGO$ are still strongly significant for 4 out of the 5 proxies, confirming that investors' reference-dependent preferences for volatility does not appear 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 give a better

chance to break 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 among firms with low nominal prices. To save space, these results are only reported in the appendix (Table A4). The stronger effect among firms with lower institutional holdings is consistent with the limits-to-arbitrage effect (e.g., Nagel, 2005). Moreover, previous studies find a positive relationship between stock prices and the institutional ownership, suggesting that individual investors prefer low-price stocks (e.g., Falkenstein, 1996, Gompers and Metrick, 2001, and Kumar, 2009). 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 since the reference-dependent preference might be a better description of individuals' risk attitudes rather than institutional investors' risk attitudes.

Lastly, Stambaugh, Yu, and Yuan (2012) show that many anomalies are stronger following high sentiment periods when more noise traders participate in the market. Indeed, Table A5 confirms that the negative lottery spread among low-CGO firms are much more significant following a high sentiment than following a low sentiment. In addition, the role of RDP in the lottery-related anomalies is more significant following a high sentiment than a low sentiment. Indeed, during high sentiment periods, the differences of the lottery spread among high-CGO firms 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 of the lottery spread among high-CGO firms 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.

4   Conclusion

In this paper, we document that the return spreads between lottery-like assets and non-lottery-like assets varies substantially across portfolios with different levels of capital gains/losses. More specifically, the previously documented underperformance of lottery-like assets are 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 asset is likely the most plausible one. In particular, the break-even effect and the aversion to loss realization suggest that following losses, investors often take up the chances 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 appear especially attractive since they provide a better chance to break even. Combined with mental accounting, 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 these anomalies. 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.

References

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

An, Li, forthcoming, Asset pricing when traders sell extreme winners and losers, *Review of Financial Studies*.

Anderson, Anne-Marie, Edward A. Dyl, 2005, Market structure and trading volume, *Journal of Financial Research* 28, 115–131.

Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva, 2013, Noisy prices and inference regarding returns, *Journal of Finance* 68, 665–713.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2013, Anomalies and financial distress, *Journal of Financial Economics* 108, 139–159.

Baker, Malcolm, Xin Pan, and Jeffrey Wurgler, 2012, The effect of reference point prices on mergers and acquisitions, *Journal of Financial Economics* 106, 49–71.

Bali, Turan, Stephen Brown, Scott Murray, and Yi Tang, 2014, Betting against beta or demand for lottery, *Working Paper*, Georgetown University.

Bali, Turan and Nusret Cakici, 2008, Idiosyncratic volatility and the cross section of expected returns, *Journal of Financial and Quantitative Analysis* 43, 29–58.

Bali, Turan, Nusret Cakici, Xuemin Yan, and Zhe Zhang, 2005, Does idiosyncratic risk really matter? *Journal of Finance* 60, 905–929.

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

Bali, Turan and Scott Murray, 2013, Does risk-neutral skewness predict the cross-section of equity option portfolio returns? *Journal of Financial and Quantitative Analysis* 48, 1145–1171.

Barber, Brad, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 56, 773–806.

Barber, Brad, and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment,*Quarterly Journal of Economics* 116, 261–292.

Barber, Brad, and Terrance Odean, 2002, Online investors: Do the slow die first, *Review of Financial Studies* 15, 455–489.

Barberis, Nicholas, and Ming Huang, 2001, Mental accounting, loss aversion, and individual stock returns, *Journal of Finance* 56, 1247–1292.

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

Barberis, Nicholas, Ming Huang, and Tano Santos, 2001, Prospect theory and asset prices, *Quarterly Journal of Economics* 116, 1–53.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2014, Prospect theory and stock returns: An empirical test, Fordham University Schools of Business Research Paper No. 2528149.

Barberis, Nicholas, and Wei Xiong, 2009, What drives the disposition effect? An analysis of a long-standing preference-based explanation, *Journal of Finance* 64, 751–784.

Barberis, Nicholas, and Wei Xiong, 2012, Realization utility, *Journal of Financial Economics* 104, 251–271.

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch, 2014, Labor hiring, investment and stock return predictability in the cross section, *Journal of Political Economy* 122, 129–177.

Benartzi, Shlomo and Richard H. Thaler, 1995, Myopic loss aversion and the equity premium puzzle *Quarterly Journal of Economics* 110, 73–92.

Ben-David, Itzhak, and David Hirshleifer, 2012, Are investors really reluctant to realized their losses? Trading responses to past returns and the disposition effect, *Review of Financial Studies* 25, 2485–2532.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169–202.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Conrad, Jennifer, Robert Dittmar, and Eric Ghysels, 2013, Ante skewness and expected stock returns, *Journal of Finance* 68, 85–124.

Conrad, Jennifer, Nishad Kapadia and Yuhang Xing, 2014, Death and jackpot: Why do individual investors hold overpriced stocks?, *Journal of Financial Economics* 113, 455–475.

Coval, Joshua, and Tyler Shumway, 2005, Do behavioral biases affect prices? *Journal of Finance* 60, 1–34.

Dhar, Ravi, and Ning Zhu, 2006, Up close and personal: Investor sophistication and the disposition effect, *Management Science* 52, 726–740.

Dougal, Casey, Joseph Engelberg, Christopher Parsons, and Edward Van Wesep, 2015, Anchoring on credit spreads, *Journal of Finance* 70, 1039-1080.

Falkenstein, Eric G., 1996, Preferences for stock characteristics as revealed by mutual fund portfolio holdings, *Journal of Finance* 51, 111-135.

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

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

Frazzini, Andrea, 2006, The disposition effect and underreaction to news, *Journal of Finance* 61, 2017–2046.

Fredertick, Shane and George Loewenstein, 1999, Hedonic adaptation, in D. Kahneman, E. Diener, and N. Schwartz [eds] *Scientific Perspectives on Enjoyment, Suffering, and Well-Being.* New York. Russell Sage Foundation.

Gao, Pengjie, Christopher A. Parsons, and Jianfeng Shen, 2014, The global relation between financial distress and equity returns, *Working Paper*, Notre Dame, UCSD and UNSW.

Garlappi, Lorenzo and Hong Yan, 2011, Financial distress and the cross-section of equity returns, *Journal of Finance* 66, 789–822.

Garlappi, Lorenzo, Tao Shu, and Hong Yan, 2008, Default risk, shareholder advantage, and stock returns, *Review of Financial Studies* 21, 2743–2778.

George, Thomas J. and Chuan-Yang Hwang, 2004, The 52-week high and momentum investing, *Journal of Finance* 59, 2145–2176.

Gompers, Paul A., and Andrew Metrick, 2001, Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229-259.

Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311–339.

Grinblatt, Mark, and Matti Keloharju, 2001, What makes investors trade? *Journal of Finance* 56, 589–616.

Ingersoll, Jonathan E. Jr. and Lawrence J. Jin, 2013, Realization utility with reference-dependent preferences, *Review of Financial Studies* 26, 723–767.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263–291.

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

Li, Jun and Jianfeng Yu, 2012, Investor attention, psychological anchors, and stock return predictability, *Journal of Financial Economics* 104, 401–419.

Locke, Peter, and Steven Mann, 2000, Do professional traders exhibit loss realization

aversion? *Working paper*, The George Washington University and Texas Christian University.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

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

Odean, Terrance, 1998, Are investors reluctant to realize their losses? *Journal of Finance* 53, 1775–1798.

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research*, 18, 109-131.

Rayo, Luis, and Gary S. Becker, 2007, Evolutionary efficiency and happiness, *Journal of Political Economy* 115, 302–337.

Shapira, Zur, and Itzhak Venezia, 2001, Patterns of behavior of professionally managed and independent investors, *Journal of Banking and Finance* 25, 1573–1587.

Shefrin, Hersh, and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance* 40, 777–791.

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

Thaler, Richard, 1980, Toward a positive theory of consumer choice, *Journal of Economic Behavior and Organization* 1, 39–60.

Thaler, Richard, 1985, Mental accounting and consumer choice, *Marketing Science* 4, 199–214.

Thaler, Richard, and Eric Johnson, 1990, Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice, *Marketing Science* 36, 613–660.

Wang, Huijun, Jinghua Yan, and Jianfeng Yu, 2014, Reference-dependent preferences and the risk-return trade-off, *Working Paper*, University of Delaware.

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

Xing, Yuhang, Xiaoyan Zhang, and Rui Zhao, 2010, What does individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis* 45, 641–662..

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

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

Zipf, George K., 1949, Human behavior and the principle of least effort. Addison-Wesley. Cambridge, MA, 498–500.

Table 1: Summary Statistics

Panel A reports the time-series averages of the monthly value-weighted excess returns (Ret^e), the intercepts of the Fama-French three-factor regression (α_{FF3}), and equal-weighted firm characteristics for five portfolios sorted by capital gains overhang (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: Grinblatt and Han's (2005) CGO 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. Frazzini's (2006) CGO 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, and 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 one month gap, $Ret_{-36,-12}$ is the cumulative return over the past three years with one year gap, RetVol is return volatility of the monthly returns over the past five years, Turnover is calculated as monthly trading volume divided by number of shares outstanding, where the volume is the reported value from CRSP for NYSE/AMEX stocks, and 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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 Grinblatt and Han's CGO, Maxret, and Oscorep, from January 1972 to December 2014 for Jackpotp and Deathp, from January 1980 to October 2014 for Frazzini's CGO, and from January 1988 to December 2014 for Skewexp. Monthly excess returns and Fama French three-factor alphas are reported in percentages. The t-statistics are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980) for returns and Newey-West (1987) adjusted standard errors with lag=36 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.

Panel A: VW 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
Grinblatt and Han's CGO										
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-stat	(1.01)	(2.06)	(14.77)	(1.87)	(0.22)	(12.68)	(18.76)	(2.55)	(-0.93)	(-2.09)
Frazzini's CGO										
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-stat	(0.38)	(1.99)	(16.43)	(4.30)	(0.96)	(13.64)	(11.70)	(-0.20)	(-2.21)	(-2.56)

Panel B: VW Returns and EW Ex-post Skewness for Five Lottery Portfolios

Proxy =	Maxret			Jackpotp			Skewexp			Deathp			Oscorep		
	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-stat	(-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)

Table 2: Double-Sorted Portfolio Returns by Grinblatt and Han's CGO and Lottery Proxies

At the beginning of every month, we independently sort stocks into five groups based on lagged Grinblatt and Han's (2005) CGO and five groups based on lagged lottery proxies. The portfolio is then held for one month. The monthly value-weighted excess returns (Panel A), and the intercepts of the Fama-French three-factor regression (Panel B), are calculated. Grinblatt and Han's CGO 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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 difference, to save space. Excess returns and FF3 alphas 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).

Panel A: Excess Return

Proxy =	Maxret				Jackpotp				Skewexp			
	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-stat	(-5.35)	(-2.31)	(2.30)	(7.50)	(-4.15)	(-2.16)	(2.30)	(7.36)	(-2.29)	(-2.74)	(-0.22)	(2.23)

Proxy =	Deathp				Oscorep							
	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-stat	(-3.04)	(0.16)	(0.85)	(3.77)	(-2.81)	(-0.48)	(2.99)	(4.70)				

Panel B: FF3 alpha

Proxy =	Maxret				Jackpotp				Skewexp			
	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-stat	(-8.36)	(-4.53)	(1.92)	(8.17)	(-7.63)	(-4.42)	(2.32)	(7.45)	(-3.59)	(-3.99)	(-1.09)	(2.52)

Proxy =	Deathp				Oscorep							
	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-stat	(-5.98)	(-1.99)	(-0.83)	(4.36)	(-6.25)	(-2.35)	(1.55)	(5.90)				

Table 3: Double-Sorted Portfolio Returns by Frazzini's CGO and Lottery Proxies

At the beginning of every month, we independently sort stocks into five groups based on lagged Frazzini's (2006) CGO and five groups based on lagged lottery proxies. The portfolio is then held for one month. The monthly value-weighted excess returns (Panel A), and the intercepts of the Fama-French three-factor regression (Panel B), are calculated. Frazzini's CGO 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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 difference, to save space. Excess returns and FF3 alphas 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).

Panel A: Excess Return

Proxy =	Maxret				Jackpotp				Skewexp			
	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-stat	(-4.70)	(-2.23)	(0.83)	(5.99)	(-3.45)	(-2.23)	(0.29)	(4.09)	(-1.88)	(-1.75)	(-0.70)	(1.55)

Proxy =	Deathp				Oscorep			
	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-stat	(-4.06)	(-0.90)	(-0.94)	(3.10)	(-2.94)	(0.85)	(-0.40)	(2.38)

Panel B: FF3 alpha

Proxy =	Maxret				Jackpotp				Skewexp			
	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-stat	(-7.56)	(-4.26)	(0.61)	(7.29)	(-6.59)	(-4.76)	(-0.30)	(4.36)	(-3.12)	(-3.07)	(-1.63)	(1.98)

Proxy =	Deathp				Oscorep			
	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-stat	(-7.12)	(-2.92)	(-2.36)	(3.76)	(-4.71)	(-0.19)	(-1.32)	(3.16)

Table 4: Lottery Spread and Raw/Residual CGO

This table reports the Fama French three-factor alphas for lottery spread (difference between top and quintile lottery portfolios) of the bottom and top quintile CGO portfolios, and their difference. 25 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 Grinblatt and Han's (2005) CGO (CGO^{GH}), Frazzini's (2006) CGO (CGO^{FR}), and residual CGO, $RCGO^{GH}$ and $RCGO^{FR}$ corresponding to CGO^{GH} and CGO^{FR}, respectively. RCGO is the residuals obtained by regressing cross-sectionally the raw CGO on previous 12- and 36-month returns, the previous 12-month average turnover, and 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, Jackpot 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), Oscorep is the predicted bankruptcy probability in the last month from Ohlson (1980). FF3 alphas 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 October 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).

FF3 alpha 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)	(-2.3)	(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)	(-1.1)	(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)

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

This table reports the Fama French three-factor monthly alphas (in percentage) for lottery spread (difference between top and quintile lottery portfolios) among the bottom and top quintile CGO portfolios, and their difference, for five double sorts robustness tests. The 25 portfolios are constructed at the end of every month from independent sorts by Grinblatt and Han's (2005) CGO and each one of five lottery proxies in tests (I) and (II). The equal-weighted and lag-gross-return-weighted portfolio alphas 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. Grinblatt and Han's CGO 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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, from January 1972 to December 2014 for Jackpotp and Deathp, and from January 1988 to December 2014 for Skewexp. The t-statistics are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Proxy	(I) Equal-Weighted			(II) Lag-Ret-Weighted			(III) Conditional Sort		
	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)	(-14.7)	(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)	(3.8)	(-8.43)	(-2.95)	(3.79)	(-7.10)	(-2.3)	(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)	(-6.3)	(1.95)	(6.32)

Table 6: Fama-MacBeth Regressions Using Grinblatt and Han's CGO

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). Turnover is calculated as monthly trading volume divided by number of shares outstanding. 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, RetVol is return volatility of the monthly returns over the past five years. VNSP is a measure of the V-shaped disposition effect calculated based on An (2013). Maxret is the maximum daily return over the past month, Jackpotp is the predicted jackpot probability from Boyer et al.(2009), Deathp is the predicted failure probability from Campbell et al.(2008), 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 are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980).

	Benchmark	Proxy = Maxret				Proxy = Jackpotp			
	(0)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
CGO	0.005	-0.012	-0.014	-0.013	-0.014	-0.008	-0.008	-0.007	-0.007
	(5.23)	(-9.06)	(-10.34)	(-7.98)	(-8.65)	(-4.77)	(-4.91)	(-3.63)	(-3.59)
Proxy		-0.034	-0.021	-0.049	-0.045	-0.392	-0.383	-0.590	-0.573
		(-3.71)	(-2.12)	(-3.35)	(-3.06)	(-6.06)	(-5.57)	(-6.51)	(-6.19)
Proxy× CGO		0.282	0.319	0.304	0.323	1.122	1.133	1.015	0.993
		(12.94)	(13.19)	(10.69)	(11.21)	(8.78)	(8.22)	(5.91)	(5.78)
Proxy×$Ret_{-12,-1}$			-0.054		-0.054		0.084		0.068
			(-2.33)		(-2.15)		(0.72)		(0.57)
Proxy× VNSP				0.160	0.228			1.257	1.238
				(1.93)	(2.55)			(2.81)	(2.70)
Ret_{-1}	-0.062	-0.056	-0.056	-0.059	-0.059	-0.051	-0.051	-0.055	-0.055
	(-15.80)	(-13.45)	(-13.54)	(-14.16)	(-14.24)	(-12.28)	(-12.39)	(-12.98)	(-13.06)
$Ret_{-12,-1}$	0.008	0.008	0.011	0.006	0.009	0.008	0.007	0.006	0.005
	(5.40)	(5.4)	(6.4)	(4.06)	(5.16)	(4.93)	(3.21)	(3.78)	(2.39)
$Ret_{-36,-12}$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(-1.83)	(-1.76)	(-1.82)	(-2.38)	(-2.39)	(-2.20)	(-2.21)	(-2.58)	(-2.61)
LOGME	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
	(-2.34)	(-2.77)	(-2.74)	(-2.88)	(-2.86)	(-5.01)	(-4.93)	(-5.47)	(-5.39)
LOGBM	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(2.36)	(2.19)	(2.19)	(2.31)	(2.31)	(1.99)	(2.00)	(2.00)	(2.01)
VNSP				0.011	0.008			0.005	0.005
				(2.24)	(1.48)			(0.94)	(0.96)
RetVol	0.003	0.014	0.013	0.003	0.003	-0.001	-0.001	-0.011	-0.011
	(0.23)	(1.01)	(0.94)	(0.24)	(0.19)	(-0.06)	(-0.06)	(-0.79)	(-0.8)
Turnover	-0.033	-0.027	-0.026	-0.024	-0.023	-0.020	-0.019	-0.016	-0.016
	(-2.32)	(-1.92)	(-1.84)	(-1.69)	(-1.60)	(-1.40)	(-1.35)	(-1.15)	(-1.12)

36

	Proxy = Skewexp				Proxy = Deathp				Proxy = Oscorep			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
CGO	-0.006	-0.004	-0.008	-0.007	0.000	0.000	-0.002	-0.002	0.001	0.001	0.002	0.002
	(-2.65)	(-1.92)	(-3.11)	(-2.56)	(0.00)	(-0.28)	(-1.41)	(-1.44)	(0.92)	(1.00)	(1.41)	(1.49)
Proxy	-0.002	-0.003	-0.003	-0.003	-10.903	-10.850	-10.432	-10.638	-0.012	-0.003	-0.047	-0.039
	(-1.72)	(-2.23)	(-2.07)	(-2.12)	(-9.30)	(-9.09)	(-6.51)	(-6.46)	(-0.94)	(-0.22)	(-2.46)	(-1.99)
Proxy× CGO	0.016	0.013	0.018	0.015	6.061	7.278	8.296	8.483	0.200	0.208	0.169	0.170
	(7.04)	(5.39)	(6.05)	(5.07)	(1.96)	(2.26)	(3.03)	(2.88)	(7.27)	(6.05)	(5.15)	(4.49)
Proxy× $Ret_{-12,-1}$		0.005		0.006		-2.546		0.329		0.014		0.003
		(2.39)		(2.69)		(-0.89)		(0.13)		(0.35)		(0.07)
Proxy× VNSP			0.004	0.000			0.111	-1.149			0.262	0.291
			(0.48)	(-0.02)			(0.01)	(-0.13)			(2.15)	(2.26)
Ret_{-1}	-0.035	-0.035	-0.038	-0.038	-0.048	-0.048	-0.065	-0.066	-0.059	-0.059	-0.062	-0.063
	(-7.25)	(-7.28)	(-7.75)	(-7.8)	(-15.18)	(-15.23)	(-16.06)	(-16.11)	(-15.04)	(-15.12)	(-15.79)	(-15.87)
$Ret_{-12,-1}$	0.006	0.004	0.005	0.002	0.006	0.007	0.004	0.004	0.008	0.008	0.006	0.006
	(3.19)	(1.51)	(2.51)	(0.69)	(3.97)	(4.33)	(2.57)	(2.22)	(5.40)	(5.05)	(4.11)	(3.85)
$Ret_{-36,-12}$	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(-2.27)	(-2.15)	(-2.46)	(-2.36)	(-2.94)	(-2.93)	(-3.35)	(-3.49)	(-2.63)	(-2.60)	(-3.17)	(-3.15)
LOGME	0.000	0.000	0.000	0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-1.10)	(-1.19)	(-1.15)	(-1.19)	(-2.80)	(-2.80)	(-2.79)	(-2.70)	(-3.27)	(-3.21)	(-3.36)	(-3.30)
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
	(0.07)	(0.07)	(0.10)	(0.10)	(4.14)	(4.15)	(4.26)	(4.09)	(1.64)	(1.66)	(1.73)	(1.76)
VNSP			0.014	0.017			0.021	0.023			0.019	0.019
			(2.21)	(2.56)			(4.77)	(5.02)			(5.31)	(5.24)
RetVol	0.004	0.004	-0.004	-0.004	0.003	0.003	-0.011	-0.014	0.000	-0.001	-0.013	-0.013
	(0.23)	(0.25)	(-0.27)	(-0.25)	(0.2)	(0.2)	(-0.77)	(-0.98)	(-0.03)	(-0.05)	(-0.89)	(-0.91)
Turnover	-0.017	-0.017	-0.011	-0.011	-0.019	-0.019	-0.012	-0.007	-0.032	-0.032	-0.029	-0.028
	(-1.62)	(-1.63)	(-0.98)	(-0.98)	(-1.33)	(-1.31)	(-0.86)	(-0.46)	(-2.23)	(-2.18)	(-1.96)	(-1.90)

Table 7: Fama-MacBeth Regressions, Robustness Checks

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 test (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 Amihud's (2002) illiquidity measure) in test (III). Variable definitions and sample period are the same as Table 7. The intercept of the regression is not reported. The t-statistics are in parentheses and are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Proxy =	(I) WLS					(II) Excluding NASDAQ stocks					(III) Excluding top illiquid decile				
	Maxret	Jackpotp	Skewexp	Deathp	Oscorep	Maxret	Jackpotp	Skewexp	Deathp	Oscorep	Maxret	Jackpotp	Skewexp	Deathp	Oscorep
CGO	-0.013	-0.005	-0.008	-0.007	-0.004	-0.016	-0.007	-0.009	-0.005	0.000	-0.014	-0.007	-0.008	-0.006	0.000
	(-6.69)	(-2.63)	(-2.63)	(-3.85)	(-2.54)	(-7.08)	(-2.55)	(-2.24)	(-2.04)	(-0.3)	(-7.47)	(-2.95)	(-2.62)	(-3.02)	(0.23)
Proxy	-0.027	-0.871	-0.002	-5.349	-0.105	-0.041	-0.580	-0.001	-9.722	-0.024	-0.058	-0.920	-0.005	-13.262	-0.084
	(-1.30)	(-5.17)	(-0.82)	(-1.94)	(-2.91)	(-2.20)	(-4.52)	(-0.38)	(-4.89)	(-1.08)	(-3.36)	(-7.48)	(-2.88)	(-6.83)	(-3.24)
Proxy×CGO	0.245	0.713	0.013	12.668	0.182	0.369	1.182	0.015	13.499	0.272	0.295	1.067	0.016	14.147	0.180
	(6.33)	(2.81)	(2.85)	(2.60)	(3.30)	(8.03)	(3.51)	(2.90)	(2.46)	(4.72)	(8.68)	(4.44)	(3.75)	(3.67)	(3.49)
Proxy×$Ret_{-12,-1}$	0.012	0.015	-0.004	-0.506	0.010	-0.004	0.317	0.005	6.688	0.008	-0.040	-0.079	0.003	-0.807	0.003
	(0.35)	(0.08)	(-1.22)	(-0.12)	(0.17)	(-0.11)	(1.72)	(1.54)	(1.96)	(0.17)	(-1.51)	(-0.53)	(1.04)	(-0.26)	(0.06)
Proxy×VNSP	0.247	3.110	-0.001	3.949	0.559	0.129	1.345	-0.001	-4.473	0.253	0.409	3.430	0.015	28.750	0.583
	(2.10)	(4.58)	(-0.07)	(0.26)	(3.07)	(1.06)	(1.87)	(-0.05)	(-0.35)	(1.59)	(3.84)	(5.68)	(1.49)	(2.38)	(3.64)
Ret_{-1}	-0.047	-0.037	-0.020	-0.043	-0.045	-0.050	-0.048	-0.022	-0.055	-0.057	-0.057	-0.054	-0.034	-0.061	-0.060
	(-8.94)	(-6.80)	(-3.20)	(-8.38)	(-9.22)	(-10.97)	(-10.2)	(-3.97)	(-12.4)	(-13.16)	(-13.00)	(-12.13)	(-6.63)	(-14.48)	(-14.5)
$Ret_{-12,-1}$	0.008	0.007	0.007	0.006	0.007	0.007	0.004	0.004	0.002	0.006	0.009	0.006	0.004	0.004	0.006
	(2.97)	(2.69)	(2.60)	(2.73)	(3.72)	(3.53)	(1.63)	(1.39)	(0.85)	(3.36)	(4.48)	(2.63)	(1.65)	(2.39)	(3.76)
$Ret_{-36,-12}$	0.000	-0.001	0.000	-0.001	-0.001	-0.001	-0.001	0.000	-0.002	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002
	(-0.44)	(-0.68)	(-0.37)	(-1.10)	(-0.90)	(-1.01)	(-1.01)	(-0.42)	(-1.94)	(-1.49)	(-2.03)	(-2.26)	(-2.15)	(-3.29)	(-2.68)
LOGME	-0.001	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	0.000	-0.001	-0.001	-0.001	-0.002	0.000	-0.001	-0.001
	(-3.4)	(-4.42)	(-1.65)	(-2.96)	(-4.15)	(-2.88)	(-4.45)	(-0.9)	(-2.38)	(-3.03)	(-3.07)	(-5.72)	(-1.2)	(-2.85)	(-3.53)
LOGBM	0.000	0.001	0.000	0.001	0.000	0.001	0.001	0.000	0.002	0.001	0.001	0.001	0.000	0.002	0.001
	(0.74)	(1.40)	(-0.36)	(2.25)	(0.61)	(2.61)	(2.78)	(0.55)	(4.62)	(2.51)	(1.89)	(1.78)	(-0.09)	(3.67)	(1.54)
VNSP	-0.005	-0.010	-0.005	0.005	0.002	0.005	-0.005	-0.008	0.013	0.004	0.001	-0.007	0.007	0.015	0.016
	(-0.74)	(-1.78)	(-0.57)	(0.83)	(0.41)	(0.78)	(-0.82)	(-0.82)	(2.00)	(0.85)	(0.20)	(-1.29)	(0.98)	(2.84)	(4.11)
RetVol	-0.008	-0.006	0.004	-0.015	-0.021	0.006	-0.004	0.002	-0.008	-0.010	0.000	-0.009	-0.004	-0.013	-0.016
	(-0.50)	(-0.31)	(0.21)	(-0.90)	(-1.20)	(0.4)	(-0.27)	(0.08)	(-0.51)	(-0.66)	(0.02)	(-0.61)	(-0.23)	(-0.89)	(-1.03)
Turnover	-0.009	-0.008	0.002	-0.004	-0.004	-0.027	-0.021	-0.017	-0.009	-0.033	-0.027	-0.020	-0.010	-0.015	-0.031
	(-0.54)	(-0.55)	(0.17)	(-0.26)	(-0.19)	(-1.93)	(-1.53)	(-1.62)	(-0.65)	(-2.27)	(-1.99)	(-1.52)	(-0.99)	(-1.09)	(-2.14)

38

Table 8: Residual Lottery Spread

This table reports the value-weighted excess returns and Fama French three-factor alphas for residual lottery spread (difference between top and quintile residual lottery portfolios) of the bottom and top quintile CGO portfolios, and their difference. 25 portfolios are constructed at the end of every month from independent sorts by Grimblatt and Han's (2005) CGO and each of the five residual lottery proxies, which are the residuals obtained by regressing cross-sectionally each of the five lottery proxies on return volatility defined as the standard deviation of the monthly returns over the past five years. 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), Oscorep is the predicted bankruptcy probability in the last month from Ohlson (1980). Returns and FF3 alphas 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).

Proxy =	Maxret			Jackpotp			Skewexp			Deathp			Oscorep		
	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1
P1	0.979	0.947		0.525	0.788		0.935	1.182		0.824	0.874		0.299	1.065	
P2	0.962	0.561		0.789	0.713		0.815	0.743		0.685	0.671		0.612	0.659	
P3	0.451	0.749		0.555	0.759		0.516	0.833		0.684	0.691		0.772	0.754	
P4	0.564	0.583		0.440	1.185		0.719	0.799		0.573	0.812		0.812	0.675	
P5	0.058	1.049		-0.150	1.323		0.257	1.131		0.227	1.047		0.347	1.085	
P5-P1	-0.921	0.102	1.023	-0.676	0.535	1.211	-0.678	-0.051	0.627	-0.597	0.173	0.770	0.048	0.021	-0.027
t-stat	(-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	-1.016	-0.023	0.993	-0.911	0.279	1.190	-0.802	-0.168	0.634	-1.190	-0.187	1.003	-0.203	-0.154	0.049
t-stat	(-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)

Table 9: Fama-MacBeth Regressions, Controlling for the Interaction Between Volatility and CGO

This table reports the time-series average of the regression coefficients from Fama-MacBeth regressions controlling for the interaction effect of return volatility and CGO. Variable definitions and sample period are the same as Table 7. The intercept of the regression is not reported. The t-statistics are in parentheses and are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Proxy =	Maxret	Jackpotp	Skewexp	Deathp	Oscorep
CGO	-0.019	-0.014	-0.014	-0.015	-0.014
	(-8.89)	(-6.06)	(-4.81)	(-6.79)	(-6.55)
Proxy	-0.040	-0.556	-0.003	-11.413	-0.039
	(-2.81)	(-6.01)	(-2.02)	(-6.95)	(-2.01)
Proxy$\times$ CGO	0.282	0.583	0.010	2.961	0.110
	(9.19)	(3.01)	(2.97)	(1.01)	(2.82)
Proxy$\times Ret_{-12,-1}$	-0.053	0.088	0.006	1.075	0.008
	(-2.12)	(0.73)	(2.90)	(0.41)	(0.19)
Proxy$\times$ VNSP	0.225	1.226	0.001	2.312	0.267
	(2.53)	(2.68)	(0.12)	(0.26)	(2.05)
RetVol$\times$ CGO	0.058	0.093	0.082	0.128	0.141
	(3.00)	(4.42)	(3.45)	(6.85)	(7.06)
Ret_{-1}	-0.060	-0.056	-0.039	-0.067	-0.064
	(-14.49)	(-13.29)	(-8.27)	(-16.58)	(-16.32)
$Ret_{-12,-1}$	0.009	0.004	0.001	0.003	0.005
	(4.98)	(2.16)	(0.39)	(1.63)	(3.32)
$Ret_{-36,-12}$	-0.002	-0.002	-0.002	-0.002	-0.002
	(-2.43)	(-2.61)	(-2.32)	(-3.43)	(-3.14)
LOGME	-0.001	-0.002	0.000	-0.001	-0.001
	(-2.94)	(-5.35)	(-1.20)	(-2.70)	(-3.24)
LOGBM	0.001	0.001	0.000	0.002	0.001
	(2.32)	(1.99)	(0.1)	(3.96)	(1.64)
VNSP	0.009	0.005	0.018	0.022	0.021
	(1.81)	(1.06)	(2.67)	(4.92)	(5.99)
RetVol	-0.006	-0.012	-0.012	-0.015	-0.011
	(-0.41)	(-0.85)	(-0.73)	(-1.07)	(-0.72)
Turnover	-0.022	-0.017	-0.009	-0.007	-0.028
	(-1.58)	(-1.21)	(-0.85)	(-0.53)	(-1.89)

Appendix I: Additional Robustness Checks

Table A1: Propensity to Sell Lottery Stocks, Individual Investors

The table presents results from probit regressions in which the dependent variable is a dummy equal to 1 if a stock was sold, and 0 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), 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 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 1 if the return since purchase is positive (zero), 0 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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.

| | 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+ x Proxy	0.0038*	0.0615**	0.0013**	0.9305**	0.0049**
	(2.10)	(5.38)	(5.78)	(7.62)	(5.55)
Ret- x 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)
Obs	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

Table A2: Propensity to Sell Lottery Stocks, Mutual Funds

The table presents results from probit regressions in which the dependent variable is a dummy equal to 1 if a stock was sold, and 0 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 to 2013. Observations are at fund-stock-report day level, where funds typically report their holdings at quarterly frequency. Following this literature, we assume trading happens on the report date. Ret^+ (Ret^-) is the return since purchase if the return since purchase is positive (negative), 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 1 if the return since purchase is positive (zero), 0 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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.

	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+ x Proxy	0.1531	-1.4538	0.0270**	11.3785**	0.1960**
	(1.79)	(-1.83)	(2.72)	(7.38)	(3.40)
Ret- x 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)
Obs	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

Table A3: Double Sorts, Robustness Checks

This table reports the Fama French three-factor monthly alphas (in percentage) for lottery spread (difference between top and quintile lottery portfolios) among the bottom and top quintile CGO portfolios, and their difference, for five double sorts robustness tests. The 25 portfolios are constructed at the end of every month from independent sorts by Grinblatt and Han's (2005) CGO and each one of five lottery proxies in tests (I) and (II). NASDAQ stocks and the top decile illiquid stocks (using Amihud's (2002) illiquidity measure) are excluded in Panels (I) and (II), respectively. The portfolio is then held for one month. Grinblatt and Han's CGO 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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, from January 1972 to December 2014 for Jackpotp and Deathp, and from January 1988 to December 2014 for Skewexp. The t-statistics are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Proxy	(I) Excluding NASDAQ Stocks			(II) Excluding Top Illiquid Decile		
	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1
Maxret	-1.62	0.39	2.01	-1.74	0.25	1.99
	(-7.63)	(2.21)	(8.17)	(-8.00)	(1.36)	(7.76)
Jackpotp	-1.32	0.41	1.73	-1.72	0.34	2.06
	(-6.74)	(2.14)	(6.55)	(-8.1)	(1.81)	(7.19)
Skewexp	-1.08	-0.28	0.81	-1.14	-0.28	0.86
	(-3.51)	(-1.22)	(2.36)	(-4.24)	(-1.13)	(2.70)
Deathp	-1.35	-0.12	1.23	-1.98	-0.44	1.54
	(-5.13)	(-0.50)	(3.97)	(-7.10)	(-2.30)	(4.72)
Oscorep	-0.96	0.19	1.15	-1.24	0.30	1.54
	(-5.15)	(1.26)	(4.96)	(-6.3)	(1.95)	(6.32)

Table A4: Double Sorts in Subsamples of Top and Bottom Institutional Ownership or Nominal Stock Price

This table reports the Fama French three-factor monthly alphas (in percentage) for lottery spread (difference between top and bottom quintile lottery portfolios) among the bottom and top tercile CGO portfolios, and their difference, within top 25% institutional ownership (or nominal stock price) and bottom 25% institutional ownership (or nominal stock price) stocks. At the beginning of every month, we first divide stocks into 3 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 lagged Grinblatt and Han's (2005) CGO and five groups based on lagged lottery proxies. The portfolio is then held for one month. Institutional ownership (IO) is the percentage of shares held by institutions each month. Grinblatt and Han's CGO 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), 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 are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Lottery Proxy	Top 25% IO			Bottom 25% IO			Top - Bottom IO		
	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)	(-2.1)	(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)

Lottery Proxy	Top 25% Price			Bottom 25% Price			Top - Bottom Price		
	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)	(0.3)	(-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)

Table A5: Lottery Spreads in High-Sentiment vs. Low-Sentiment Periods

This table reports the benchmark-adjusted returns (in percentage) for lottery spread (difference between top and bottom quintile lottery portfolios) among the bottom and top quintile CGO portfolios following high sentiment and low sentiment periods. The 25 portfolios are constructed at the end of every month from independent sorts by Grinblatt and Han's (2005) CGO and each one of five lottery proxies. The portfolio is then held for one month. The benchmark-adjusted returns in high and low sentiment periods are estimates of a_H and a_L in the regression: $R_{i,t} = a_H d_{H,t} + a_M d_{M,t} + a_L d_{L,t} + bMKT_t + cSMB_t + dHML_t + \epsilon_{i,t}$. $d_{H,t}$, $d_{M,t}$ and $d_{L,t}$ are dummy variables indicating high, middle, and low sentiment periods, and $R_{i,t}$ is the excess percent returns in month t. A month is considered to be a high sentiment month if the value of the BW sentiment index at the end of the previous month is above the 30% percentile of the historical BW sentiment index till this month, the low sentiment months are those below the historical 30% percentile, and the rest of the months are the middle sentiment months. Grinblatt and Han's CGO 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.(2009), Deathp is the predicted failure probability in the last month from Campbell et al.(2008), Oscorep is the predicted bankruptcy probability in the last month from Ohlson (1980). The sample period is from July 1965 to January 2011 for Maxret and Oscorep, from January 1972 to January 2011 for Jackpotp and Deathp, and from January 1988 to January 2011 for Skewexp. The t-statistics are in parentheses calculated based on the heteroskedasticity-consistent standard errors of White (1980).

Lottery Proxy	High Sentiment Periods			Low Sentiment Periods			High-Low Sentiment Periods		
	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1
Maxret	-2.59	-0.09	2.50	-0.41	0.15	0.56	-2.18	-0.25	1.94
	(-5.03)	(-0.23)	(4.44)	(-0.61)	(0.26)	(0.73)	(-2.56)	(-0.35)	(2.04)
Jackpotp	-2.63	0.02	2.64	-0.69	-0.05	0.64	-1.94	0.06	2.00
	(-4.78)	(0.03)	(3.72)	(-1.34)	(-0.09)	(1.17)	(-2.61)	(0.09)	(2.26)
Skewexp	-1.57	-1.55	0.02	2.67	2.04	-0.63	-4.24	-3.59	0.65
	(-1.17)	(-2.02)	(0.02)	(2.11)	(2.73)	(-0.48)	(-2.36)	(-3.32)	(0.35)
Deathp	-2.63	-0.33	2.30	-0.07	-0.54	-0.47	-2.56	0.21	2.77
	(-3.87)	(-0.46)	(3.05)	(-0.12)	(-0.75)	(-0.56)	(-2.92)	(0.20)	(2.44)
Oscorep	-1.59	0.36	1.95	-0.97	0.00	0.97	-0.63	0.36	0.99
	(-3.95)	(1.00)	(3.68)	(-1.77)	(0.00)	(1.35)	(-0.94)	(0.64)	(1.15)

Appendix II: Definitions of Key Variables

In this appendix, we describe the details for constructing several CGO and lottery measures.

CGO^{GH}: our first CGO measure is constructed following Grinblatt and Han (2005) (Equation (9), page 319, and Equation (11), page 320). At each week t, the reference price for each individual stock is defined as

$$RP_t^{GH} = k^{-1} \sum_{n=1}^{T} \left(V_{t-n} \prod_{\tau=1}^{n-1} (1 - V_{t-n+\tau}) \right) P_{t-n},$$

where V_t is turnover in week ts, T is 260, the number of weeks in the previous five years, P_t is the stock price at the end of week t, and k is a constant that makes the weights on past prices sum to one. Weekly turnover is calculated as the weekly trading volume divided by the number of shares outstanding. To address the issue of double counting of the volume for NASDAQ stocks, we follow Anderson and Dyl (2005). They propose a rough rule of thumb to scale down the volume of NASDAQ stocks by 38% after 1997 and by 50% before 1997 to make it roughly comparable with the volume on NYSE. Further, we only include observations that have at least 200 weeks of nonmissing data in the previous five years. As argued by Grinblatt and Han (2005), the weight on P_{t-n} reflects the probability that the share purchased at week $t-n$ has not been traded since. The capital gains overhang (CGO) at week t is then defined as

$$CGO_t^{GH} = \frac{P_{t-1} - RP_t^{GH}}{P_{t-1}}.$$

To avoid market microstructure effects, the market price is lagged by one week. Finally, to obtain monthly CGO, we simply use the last week CGO within each month. Since we use five-year daily data to construct CGO and require a minimum of 150-week nonmissing values in the calculation, this CGO variable ranges from January 1965 to December 2014.

CGO^{FR}: we also adopt an alternative CGO measure using mutual fund holding data as in Frazzini (2006) (Equation (1), page 2022, and Equation (2), page 2023). At each month t, the reference price for each individual stock is defined as

$$RP_t^{FR} = \phi^{-1} \sum_{n=0}^{t} V_{t,t-n} P_{t-n},$$

where $V_{t,t-n}$ is the number of shares purchased at date $t-n$ that are still held by the original purchasers at date t, P_t is the stock price at the end of month t, and ϕ is a normalizing

constant such that $\phi = \sum_{n=0}^{t} V_{t,t-n}$. The stock price at the report date is used as a proxy for the trading price. Following Frazzini (2006), when trading, fund managers are assumed to use the "first in, first out" method to associate a quantity of shares in their portfolio to the corresponding reference price. Fund holdings are adjusted for stock splits and assumed to be public information with a one-month lag from the file date. The quarterly holdings data are merged with CRSP and filtered to eliminate potential errors in the data. The CGO at month t is then defined as the normalized difference between current price and reference price:

$$CGO_t^{FR} = \frac{P_t - RP_t^{FR}}{P_t}.$$

The resulting sample period is from April 1980 to October 2014.

Jackpotp: The predicted jackpot probability is constructed from the baseline model in Conrad, Kapadia, and Xing (2014) (Table 3, Panel A, page 461). In particular, for each firm, we first estimate the baseline logit model using data from the past 20 years at the end of June every year:

$$Prob_{t-1}(Jackpot_{i,t} = 1) = \frac{exp(a + b \times X_{i,t-1})}{1 + exp(a + b \times X_{i,t-1})},$$

where $Jackpot_{i,t}$ is a dummy that equals to 1 if firm i's log return in the next 12 month period is larger than 100%. The vector $X_{i,t-1}$ is a set of firm-specific variables known at time $t-1$, including skewness of log daily returns (centered around 0) over the last 3 months, log stock return over the past year, firm age as the number of years since appearance on CRSP, asset tangibility as the ratio of gross PPE (property plant and equipment) to total assets, the log of sales growth over the prior year, detrended stock turnover as the difference between the average past 6-month turnover and the average past 18-month turnover, volatility as the standard deviation of daily returns (centered around 0) over the past 3 months, and the log of market equity in thousands. Next, we use these estimated parameters to construct the out-of-sample predicted jackpot probability (Jackpotp). We reestimate this model for each firm every year from 1951, so our first set of out-of-sample predicted jackpot probabilities is from January 1972.

Skewexp: The expected idiosyncratic skewness is calculated in two steps following Boyer, Mitton, and Vorkink (2010) (Table 2, Model 6, page 179). First, we estimate the following cross-sectional regressions separately at the end of each month t:

$$is_{i,t} = \beta_{0,t} + \beta_{1,t} is_{i,t-60} + \beta_{2,t} iv_{i,t-60} + \lambda_t' X_{i,t-60} + \varepsilon_{i,t},$$

where $is_{i,t}$ and $iv_{i,t}$ denote the historical estimates of idiosyncratic volatility and skewness relative to the Fama and French three-factor model, respectively, for firm i using daily stock data over the past 60 months till month t. $X_{i,t}$ is a set of firm-specific variables including momentum as the cumulative returns over months $t-72$ through $t-61$, turnover as the average daily turnover in month $t-60$, small-size market capitalization dummy, medium-size market capitalization dummy, industry dummy based on the Fama-French 17-industries definition, and the NASDAQ dummy. After we have these regression parameters, the expected idiosyncratic skewness for each firm i at the end of each month t is then computed in the second step:

$$Skewexp_t \equiv E_t[is_{i,t+60}] = \beta_{0,t} + \beta_{1,t}is_{i,t} + \beta_{2,t}iv_{i,t} + \lambda_t' X_{i,t}.$$

Similar to Boyer, Mitton, and Vorkink's (2010) baseline database, our expected idiosyncratic skewness measure dates back to January 1988, due to limited availability of the NASDAQ turnover data for the earlier sample.

   Deathp: The predicted failure probability is constructed following Campbell, Hilscher, and Szilagyi (2008) (Table IV, 12 month lag, page 2913). In particular, for each firm, we first use the most recently available Compustat quarterly and CRSP data to compute a distress score:

$$Distress_t = -9.16 - 20.26NIMTAAVG_t + 1.42TLMTA_t - 7.13EXRETAVG_t$$
$$+ 1.41SIGMA_t - 0.045RSIZE - 2.13CASHMTA_t + 0.075MB_t - 0.058PRICE_t,$$

where

$$NIMTAAVG_{t-1,t-12} = \frac{1-\phi^2}{1-\phi^{12}}(NIMTA_{t-1,t-3} + \cdots + \phi^9 NIMTA_{t-10,t-12})$$
$$EXRETAVG_{t-1,t-12} = \frac{1-\phi}{1-\phi^{12}}(EXRET_{t-1} + \cdots + \phi^{11}EXRET_{t-12}),$$

The coefficient $\phi = 2^{-1/3}$ implies that the weight is halved each quarter. NIMTA is net income (Compustat quarterly item NIQ) divided by the sum of market equity and total liabilities (Compustat quarterly item LTQ). The moving average NIMTAAVG uses a longer history of losses that can better predict bankruptcy than a single month. $EXRET = log(1+R_{it}) - log(1+R_{S\&P500,t})$ is the monthly log excess return on each firm's equity relative to the S&P 500 index. To use its moving average, the model assumes that a sustained decline in stock market value could better predict bankruptcy than a sudden stock price decline in

a single month. TLMTA is total liability divided by the sum of market equity and total liabilities. SIGMA is the volatility of daily stock returns over the past three months. RSIZE is the log ratio of each firm's market equity to that of the S&P 500 index. CASHMTA is the ratio of cash and short-term investments (Compustat quarterly item CHEQ) divided by the sum of market equity and total liabilities. MB is the market-to-book equity. PRICE is the log of stock price, winsorized at \$15. The corresponding distress probability is then calculated as:

$$Deathp_t = \frac{1}{1 + exp(-Distress_t)}.$$

This measure requires to use Compustat quarterly data, so the sample period starts from January 1972.

Oscorep: The predicted bankruptcy probability is calculated based on a set of annual accounting information following Ohlson(1980) (Table 4, Model 1, page 121). We first calculate an O-score every year for each firm as following :

$$
\begin{aligned}
Oscore_t = {}& -1.32 - 0.407 log(\frac{\text{total assets}_t}{\text{GNP price-level index}_{t-1}}) + 6.03(\frac{\text{total liabilities}_t}{\text{total assets}_t}) \\
& - 1.43(\frac{\text{working capital}_t}{\text{total assets}_t}) + 0.076(\frac{\text{current liabilities}_t}{\text{current assets}_t}) - 2.37(\frac{\text{net income}_t}{\text{total assets}_t}) \\
& - 1.72(1 \text{ if total liabilities} > \text{total assets at year t, 0 otherwise}) \\
& - 1.83(\frac{\text{funds from operations}_t}{\text{total liabilities}_t}) - 0.521(\frac{\text{net income}_t - \text{net income}_{t-1}}{|\text{net income}_t| + |\text{net income}_{t-1}|}) \\
& + 0.285(1 \text{ if net loss for the last two years, 0 otherwise}),
\end{aligned}
$$

and the corresponding bankruptcy probability is then obtained from the logistic function:

$$Oscorep = \frac{1}{1 + exp(-Oscore)}.$$