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

Li An, Huijun Wang, Jian Wang, and Jianfeng Yu

October 2015

Preliminary Version

Abstract

This paper studies the role of reference-dependent preferences in explaining several anomalies related to lottery-like assets. Recent studies find that lottery-like assets significantly underperform non-lottery-like assets. We document that lottery anomalies are significantly stronger among stocks where investors have lost money. Among stocks where investors have profited, evidence for lottery-related anomalies is either very weak or even reversed. We consider several potential explanations for this empirical fact. Our findings lend support for the reference-dependent preference under which investors are averse to losses and prefer lottery-like assets following prior losses as such assets provide a better chance to recover losses. Moreover, Our findings are robust to five different measures of the lottery feature of stocks and provide a unified framework to understand lottery-related anomalies that are associated with these measures.

JEL Classification: G02, G12, G14

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

---

*We thank Nicholas Barberis, Jeremy Page (discussant), Yu Yuan, and seminar participants at University of Minnesota, PBC School of Finance at Tsinghua University, 2015 Northern Finance Association Conference for helpful comments and discussions. We also thank Terry Odean for the brokerage data. Author affiliation/contact information: An: PBC School of Finance Tsinghua University, Office 1-518, 43 Chengfu Road, Haidian District, Beijing 100083, China. Email: anl@pbcsf.tsinghua.edu.cn, Phone: 86-10-62797840. Wang: Lerner College of Business and Economics, University of Delaware, 307A Purnell Hall, Newark, DE 19716. Email: wangh@udel.edu, Phone: 302-831-7087. Wang: Research Department, Federal Reserve Bank of Dallas, 2200 N. Pearl Street Dallas TX 75265, Email : jian.wang@dal.frb.org, Phone: 214-922-6471. Yu: Carlson School of Management, University of Minnesota, 321 19th Avenue South, Suite 3-122, Minneapolis, MN 55455. Email: jianfeng@umn.edu, Phone: 612-625-5498, Fax: 612-626-1335. All views are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

Many prior studies find that lottery-like stocks (such as stocks with large daily maximum returns, large jackpot probability, large expected skewness, and large default probability) tend to significantly underperform non-lottery-like stocks. One popular explanation is that investors have a strong preference for lottery-like assets, leading to the overpricing of these assets. There is a small probability for lottery-like assets to earn extremely high returns. Thus, the strong preference for lottery-like assets could come from the overweighting of these extremely high returns (see., Barberis and Huang (2008)). Indeed, the overweighting of low probability events is a key feature of prospect theory (PT) utility.

Motivated by prior studies on lottery-related anomalies, we use five proxies to measure the extent to which a stock exhibits lottery-like payoffs (i.e. skewness). Here, we use skewness, lottery, and lottery-like feature of a stock interchangeably. The five measures are maximum daily returns, jackpot probability, expected idiosyncratic skewness, failure probability, and bankruptcy probability. We first document a new empirical fact. Specifically, using these five skewness measures, we show that lottery-related anomalies are significantly stronger among stocks with negative capital gain overhang (CGO). Among stocks with high CGO, evidence for lottery-related anomalies is either very weak or even reversed. The reason that we study the skewness-return relation among firms with different levels of CGO is motivated by a specific conjecture that we will discuss in more detail in Section 3. We delay further discussion of this hypothesis until then. The basic idea is that investors' preference for skewness might be different depending on whether his investment is in gain or in loss relative to a reference point. Thus, by separating firms with capital-gain investors from those with capital-loss investors, we can investigate the refined skewness-return relation within each group.

Specifically, we first utilize the method in Grinblatt and Han (2005) to calculate the capital gains overhang (CGO) for individual stocks, which is essentially the normalized difference between the current stock price and the reference price. As a robustness check, we also compute an alternative measure of CGO based on mutual fund holdings as in Frazzini (2006). We then sort all individual stocks into portfolios based on lagged CGO and various measures of skewness. We find that, among firms with prior capital losses, the returns of firms with high maximum daily returns are 138 basis points (bps) lower per month than those of firms with low maximum daily returns. By sharp contrast, among firms with prior capital gains, the returns of firms with high maximum daily returns are 54 bps higher per month than those of firms with low maximum daily returns. Similar results hold when the lottery feature is measured by jackpot probability, expected idiosyncratic skewness, failure probability, and bankruptcy probability.

We then consider several possible explanations for this new empirical fact. First, we investigate the role of reference-dependent preferences (RDP) and mental accounting (MA) in these lottery-related anomalies. Under RDP, investors' risk-taking behavior in the loss region can be different from that in the gain region. For example, prospect theory (PT) posits that individuals tend to be risk seeking and also have a strong desire to break even following prior losses relative to a reference point. Lottery-like assets are particularly attractive in this case since they provide a higher chance to recover prior losses. On the other hand, when investors faces prior gains, their preference for lottery-like assets is not as strong since there is neither a break-even effect nor a need to avoid losses. Instead, due to the high volatility of lottery-like stocks, investors with mental accounting tend to dislike this type of stocks since investors are risk averse in their gain region.

The mental accounting of Thaler (1980, 1985) provides a theoretical foundation for the way in which decision makers set reference points for each asset they own. The key idea underlying mental accounting 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. As a result, if the arbitrage forces are limited, lottery-like stocks could be overvalued compared to non-lottery-like stocks among the set of stocks where investors face prior losses, leading to lower future returns for lottery-like stocks. By contrast, among the stocks where investors face capital gains, lottery feature may not be negatively associated with future returns. The correlation could even turn positive since investors with capital gains usually dislike the high volatility associated with lottery-like stocks. Thus, RDP together with MA can potentially account for the new empirical fact.

However, Barberis and Xiong (2009) show that the above static argument may not be valid in a dynamic setting. Thus, before fully embracing the above argument, it would be helpful to develop a formal model in a dynamic setting, which is beyond the scope of our study. Despite this concern, we want to emphasize that the main purpose of this paper is to show that RDP seems to play a role in the lottery-related anomalies, and our results point to the usefulness of constructing such a dynamic model in future research.

The second possible explanation for the above finding is an underreaction to news story. To see why, take the failure probability as an example. Firms with high CGO are likely to have experienced good news. If information travels slowly across investors, then firms with high CGO tend to be undervalued on average. More importantly, information may travel even slower among firms with higher failure probability since information uncertainty is likely to be higher and arbitrage forces may also be more limited among these firms. Consequently, among the firms with a high CGO, firms with a high failure probability are likely to be more undervalued than firms with a low failure probability. As a result, among firms with recent good news, a positive relationship between failure probability and return is observed. On the other hand, firms with low CGO have probably experienced negative news and therefore have been overpriced due to underreaction. Similarly, this overpricing effect should be stronger when failure probability is high. Thus, there is a negative relationship between the failure probability and future returns among firms with low CGO.

The third possible explanation is the general 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 dropped in value), firms with high CGO experience higher selling pressure and thus are underpriced. Since stocks with high skewness, especially those firms close to default, tend to have higher arbitrage costs, the mispricing effect should be stronger among lottery-type firms. Similar to the underreaction to news story, this general mispricing effect could potentially explain the negative skewness-return relation among low-CGO firms and the positve skewness-return relation among high-CGO firms. Notice that the mechanism based on RDP is different from this mispricing story, since it does not require CGO to be a proxy for mispricing. It only needs that investors' preference for skewness depends on a reference point.

To explore these possible mechanisms, we perform a series of Fama-MacBeth regressions. First, we discuss the asset pricing implications of the reference-dependent preference (RDP), in particular, the effect of RDP on lottery-related anomalies. We argue that RDP could generate the same skewness-return relation as the new empirical pattern presented in Section 2 in this paper: high-skewness firms have lower returns among firms with negative CGO, and this negative skewness-return correlation is weaker and even reversed among firms with positive CGO. This suggests that reference-dependent preferences and mental accounting jointly could help account for the lottery anomalies documented in previous literature. Second, we show that our results still hold when we control for the interaction of past returns (proxy for past news) and skewness proxies. Third, we control for a battery of additional variables such as shares turnover, idiosyncratic volatility, and a mispricing measure derived from the V-shaped disposition effect. The effect of CGO on the lottery-related anomalies remains significant. In particular, we show that the mispricing role of CGO due to the disposition effect does not drive our results. Rather, it is the strong preference for lottery-like assets to recover prior losses that drives our key results. Moreover, this effect is robust to different subsamples, such as the exclusion of NASDAQ stocks and illiquid stocks.

There are many existing theoretical and empirical studies exploring the role of reference point in asset prices. Barberis and Huang (2001), Barberis, Huang, and Santos (2001), and Barberis and Huang (2008), for example, 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.[1] 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.[2] More recently, Barberis and Xiong (2009, 2012), and Ingersoll and Jin (2013) provide theoretical models of realization utility with RDP. Wang, Yan, and Yu (2014) use RDP to understand the lack of positive risk-return trade-off in the data.

A large literature studies the relationship between our various measures of skewness and expected returns. Boyer, Mitton, and Vorkink (2010) find that expected idiosyncratic skewness and returns are negatively correlated. Bali, Cakici, and Whitelaw (2011) find that maximum daily return in the past month is negatively associated future returns.[3] 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 show that firms which have a high potential for default also tend to have a relatively high probability of extremely large returns (i.e., jackpot), and that stocks with high predicted probabilities for extremely large payoffs also earn abnormally low average returns.[4] All

[1] 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 several 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.

[2] There are several other studies which also apply the reference-dependent feature in decision making to understand a few financial phenomena. 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 (2014) on credit spread.

[3] 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.

[4] In addition, several studies have employed option data to study the relation between various skewness measures and future returns. See, e.g., Xing, Zhang, and Zhan (2010), Bali and Murray (2013), and Conrad,

of these findings are consistent with investor preferences for skewed, lottery-like payoffs. The existing studies typically suggest that the mechanism responsible for the pattern is an *unconditional* preference for skewness, which is due to overweighting of small probability events. In particular, probability weighting implies that the lottery-related anomalies are independent of reference points and thus should be equally strong among stocks where investors face losses and stocks where investors face gains. We differ from previous studies by showing that the negative skewness-return relation exists much stronger among firms with capital losses, whereas this negative relation is weak and insignificant or even reversed among firms with capital gains. This new empirical finding suggests that RDP may be an important source that underlies investor's preference for lottery, above and beyond the role of probability weighting.[5]

The rest of the paper is organized as follows. Section 2 describes the definition of skewness proxies and presents a new empirical finding we document in this paper. Section 3 discusses several possible explanations for this new empirical finding, paying special attention to RDP since it motivates our key conditional variable, CGO. Additional robustness tests are covered in Section 4. Section 5 concludes.

2 Heterogeneity in the Skewness-Return Relation: A New Empirical Fact

In this section we present a new empirical finding regarding the role of CGO on the skewness-return relationship. To proceed, we first define the key variables used in our analysis. We then report summary statistics, the double-sorting portfolio and the Fama-MacBeth regressions analysis.

2.1 Definition of Key Variables

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,

Dittmar, and Ghysels (2013). The sample periods for these studies are typically short due to the availability of option data.

[5]To clarify, our results do not counter the existence of overweighting of small probability events. In facts, we find that the negative skewness-return relation is still generally significant among stocks around zero CGO region; this evidence seems to point to an independent role for probability weighting.

and mutual fund holdings data are obtained from the Thomson Financial CDA/Spectrum Mutual Funds database. To construct stock-level variables, we begin with the data of all US common stocks traded in NYSE, AMEX, and NASDAQ from 1962 to 2014. We then require our observations 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. After applying these common filters, we further construct measures of capital gains overhang and several other variables for lottery-related anomalies as follows.

2.1.1 Capital Gains Overhang

We adopt two definitions of CGO following the previous literature. This section briefly describes the construction of these variables. More details on calculating these measures are provided in Appendix II.

CGO^{GH}: Grinblatt and Han (2005) propose a turnover-based measure to calculate the reference price and CGO (Equation (9), page 319, and Equation (11), page 320 of Grinblatt and Han (2005)). The CGO is the normalized difference between the current stock price and the reference price, and the reference price is simply a weighted average of past stock prices. The weight given to each past price is based on past turnover reflecting the probability that the share purchased at a certain date has not been traded since then. Moreover, a minimum of 150 weeks of nonmissing values over the past five years are required in the CGO calculation. Since we use five-year data to construct CGO, this CGO variable ranges from January 1965 to December 2014.

CGO^{FR}: In addition to the turnover-based measure of CGO, we also adopt an alternative measure using mutual fund holding data as in Frazzini (2006) (Equation (1), page 2022, and Equation (2), page 2023 of Frazzini (2006)). Similar to Grinblatt and Han (2005), the CGO is defined as the normalized difference between the current price and the reference price, but now the time series of net purchases by mutual fund managers and their cost basis in a stock are used to compute a weighted average reference price. The advantage of this approach is that it can exactly identify the fraction of shares purchased at a previous date that is still held by the original buyers at the current date. However, the resulting sample period is shorter, starting from April 1980 to October 2014. Also this approach assumes that mutual fund managers are a representative sample of the cross-section of shareholders.

2.1.2 Lottery Measures

We use several variables to proxy for the lottery feature based on prior studies. In particular, Boyer, Mitton, and Vorkink (2010) find that expected idiosyncratic skewness and returns are negatively correlated. Bali, Cakici, and Whitelaw (2011) find that the maximum daily return in the past month is negatively associated with future returns. Ohlson (1980) finds that firms close to bankruptcy earn lower subsequent returns. Relatedly, Campbell, Hilscher, and Szilagyi (2008) show that firms with a high probability of failure have abnormally low average future returns. Conrad, Kapaida, and Xing (2014) show that firms which have a high potential for failure also tend to have a relatively high probability of extremely large payoffs (jackpot), and that stocks with high predicted probabilities for extremely large payoffs earn abnormally low average returns. In addition, they show that the low returns of the firms with high failure probability documented by Campbell, Hilscher, and Szilagyi (2008) are concentrated in the firms in the top 30% of realized daily log return skewness over the past three months and are not present in the firms in the bottom 30% of realized past skewness. Thus, they argue that a preference for lottery-like payoffs explains why distressed stocks earn lower subsequent returns. In sum, all of these findings are consistent with an investor preference for skewed, lottery-like payoffs.

Motivated by these findings, we choose five variables to proxy for the lottery-like feature. These measures include the maximum daily return in the previous month, the probability of jackpot returns, defined as log returns greater than 100% over the next year, expected idiosyncratic skewness, defined as in Boyer, Mitton, and Vorkink (2010), failure probability defined as in Campbell, Hilscher, and Szilagyi (2008), and bankruptcy probability defined in Ohlson (1980). We briefly describe how these variables are computed and provide more details on the construction of these measures in Appendix II.

Maxret: Bali et al. (2011) document a significant and negative relation between the maximum daily return over the past month and the returns in the future. They conjecture that this negative relation is due to preferences for lottery-like stocks among investors. They also show that firms with larger maximum daily returns have higher return skewness. Therefore, we use this maximum daily return (Maxret) within each month for each stock as our first proxy.

Jackpotp: Conrad et al. (2014) show that stocks with high predicted probability of extremely large payoffs earn abnormally low subsequent returns. They suggest that this finding is consistent with an investor preference for skewed, lottery-like payoffs. Thus, we use the predicted jackpot probability from their baseline model as our second proxy (Table 3, Panel A, page 461 of Conrad et al. (2014)). Our out-of-sample predicted jackpot probabilities start from January 1972.

Skewexp: Boyer et al. (2009) estimate a cross-sectional model of expected idiosyncratic skewness and find that it negatively predicts future returns. We use the expected idiosyncratic skewness from their model as our third proxy (Table 2, Model 6, page 179 of Boyer et al. (2009)). Due to limited availability of NASDAQ turnover data in the earlier years, this measure starts from January 1988.

Deathp: Campbell et al. (2008) find that stocks with high predicted failure probability earn abysmally low subsequent returns. Since distress stocks tend to have positive skewness, they conjecture that one potential explanation could be that investors have strong preferences for positive skewness which push up the prices of such stocks and lead to lower future returns. We construct this proxy using their logit model (Table IV, 12 month lag, page 2913 of Campbell et al. (2008)) and use it as our fourth proxy. This measure requires Compustat quarterly data, so the sample period starts from January 1972.

Oscorep: Finally, Ohlson (1980) develops a model to predict a firm's probability of bankruptcy based on a set of accounting information. Following his approach, we calculate firms' predicted bankruptcy probability based on the O-score (Table 4, Model 1, page 121 of Ohlson (1980)) and use it as our fifth proxy.

2.2 Summary Statistics and One-Way Sorts

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

At the end of month t, we sort firms into quintiles based on CGO and lottery proxies. Portfolio returns are value-weighted returns in month $t+1$, while other firm characteristics 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, lower volatility, and they outperform low-CGO firms in the following month. In addition, the Fama-French three-factor alphas for the CGO portfolio spread are significant for both Grinblatt and Han's (2005) measure and Frazzini's (2006) measure. 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 based on lottery proxies, which include the max daily return, jackpot probability, expected idiosyncratic skewness, death probability, and bankruptcy probability. Consistent with previous studies on each of these anomalies, high-skewness portfolios underperform low-skewness portfolios, and the difference between high- and low-skewness portfolios is significant, especially in terms of the Fama-French three-factor alphas. Panel B also reports ex-post skewness for each portfolio, measured by the time series mean of cross-sectional average stock-level skewness calculated using daily stock returns in the next 12 months. This is to assure that our lottery proxies calculated at the portfolio formation time capture stocks' lottery feature in the future. As expected, we find ex-post skewness increases monotonically from low- to high-skewness portfolios for all five proxies.

2.3 Double Sorts

As shown in the previous subsection, our five lottery measures *unconditionally* predict future returns in a similar fashion as 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 then 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 five proxies for lottery-like feature. Panel A reports excess returns for these portfolios, while Panel B presents the Fama-French three-factor alphas. Because of the independent sorting, we have a similar spread of the skewness proxy in the high-CGO group (CGO5) and the low-CGO group (CGO1). However, we see distinct patterns for the future returns in these two groups. Take Maxret for example: following previous loss (CGO1), high-Maxret stocks *underperform* low-Maxret stocks by 1.38% per month in excess return, with the t-stat equal to -5.35. In contrast, following a previous gain (CGO5), the negative relation 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 portfolio is about -0.24% per month (from 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.

For the other four proxies, the pattern remains similar. 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, suggesting that lottery anomalies are significantly stronger among prior losers. In addition, this return pattern also holds for the Fama-French three-factor alphas, as shown in Panel B.[6] More interestingly, Panel B shows that almost all of the return spreads between low- and high-skewness firms among low-CGO firms is due to the negative alpha of the lottery-like assets. Taking Maxret as an example, the long-leg has an alpha of 0.35% per month, whereas the short-leg has an alpha of -1.76% per month. This is consistent with the notion that facing prior losses, the preference for lottery-like assets increases the demand for these assets. Due to limits to arbitrage and especially short-sale impediments, this excess demand drives up the price of lottery-like assets and leads to subsequent low returns for these assets.[7]

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 lottery-like stocks slightly outperform (in terms of excess returns) the non-lottery-like stocks by 0.54%, 0.69%, -0.05%, 0.24%, and 0.53% per month for the five proxies, respectively. Three out of these five return spreads are significant. The patterns are similar for the Fama-French 3-factor alphas, three out of five spreads being at least marginally significant and the other two negative spreads insignificant.

It is also worth noting that the lottery-like assets also underperform the non-lottery-like assets among the middle-CGO group (CGO3). These stocks are generally neither winners nor losers, but have a CGO close to zero. This suggests that besides the different effect of capital gains vs. losses on investors' preference for lottery-like assets, some other forces, such as probability weighting as proposed by previous literature, also play a role in lottery

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

[7]Relatedly, 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.

anomalies.

In addition, 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 also 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, the differences between lottery spreads among high-CGO firms and those among low-CGO firms are 1.88%, 1.26%, 0.56%, 1.10%, and 0.69% per month with t-stats 5.99, 4.09, 1.55, 3.10, and 2.38, respectively, for the five lottery-feature proxies. The sample period in Table 3 is shorter due to the availability of the mutual fund holdings data. As a result, the t-statistics are slightly lower than those in Table 2. However, the economic magnitude remains largely the same.

Using Frazzini's (2006) CGO, the lottery spreads (alphas) among high-CGO firms are very close to zero and only one of them 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 high CGO.

One caveat of using raw CGO measure is that, since CGO may correlate with other stock characteristics, in particular, past returns and shares turnover, it may be concerning that the results in Tables 2 and 3 are driven by effects other than the capital gains that investors face. To address this concern, we sort stocks based on the residual CGO after controlling for other stock characteristics. To construct the residual CGO, 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 turnover and previous 12-month return, and an interaction term between turnover and NASDAQ dummy. Table 4 reports the Fama French three-factor alphas for various portfolios. Specifically, it reports the lottery spreads (P5-P1 based on lottery proxy) among firms with high residual CGO, the lottery spread among firms with low residual CGO, and the difference between these two spreads for all five lottery proxies based on both Grinblatt and Han's (2005) and Frazzini's (2006) procedures. To facilitate comparison, we include lottery spreads based on raw CGO, as a summary of the results presented in Table 2 and Table 3. Using residual rather than raw CGO delivers results that support our hypothesis as well. Taking the residual CGO 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).

As noted by Fama and French (2008), the properties of equal-weighted returns are dominated by the behavior of very small firms. Thus, we have been focusing on value-weighted portfolio returns and we have excluded penny firms from our sample throughout our analysis. On the other hand, the properties of value-weighted returns are dominated by the behavior of a small number of very large firms because of the well-known heavy tails of the size distribution in the U.S. stock market (Zipf (1949)). Thus, in Table 5, we report both equal- and lagged-gross-return-weighted portfolio alphas.[8] The lagged-gross-return-weighted portfolio returns can further mitigate the bias due to temporary deviations of trade prices from fundamental values (Asparouhova, Bessembinder, and Kalcheva (2013)). Thus, reporting the properties of average equal-weighted returns and lagged-gross-return-weighted returns of the lottery portfolios allows us to mitigate the extreme influence of the very large firms in the pure value-weighted returns, and thus may better characterize the role of CGO for an average firm in the economy. The results in Table 5 again reveal 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 the lottery spreads are mostly positive and significant in many cases or insignificant and negative. In addition, to ensure that the pre-ranked skewness remain similar across different CGO quintiles, we have used independently double-sorted portfolios so far. If we use conditionally sorted, the pattern remains quantitatively similar. This finding is even stronger in certain cases as shown in Panel III of Table 5.[9]

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 might influence returns. However, sorting on three or more variables is impractical. Thus, to examine other possible mechanisms, we perform a

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

[9]To save space, we only report these alternative portfolio returns based on Grinblatt and Han's (2005) CGO measure. The results based on Frazzini's (2006) measure are quantitatively similar and available upon request.

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 shares turnover. Following Conard et al (2014), independent variables are winsorized at their 5th and 95th percentiles. The benchmark regression (0) in Table 6 shows that the coefficient of GCO is significant and positive, confirming the Fama-MacBeth regression results of Grinblatt and Han (2005). For each skewness proxy, regression (1) includes an interaction term between CGO and the proxy. The coefficient estimate of the interaction term is always positive and significant for all skewness proxies, confirming the simple double-sorting analysis in the last section. This is true even after we control for size, book-to-market, past returns, stock return volatility, and shares turnover. It is also noteworthy that the proxy itself typically appears to be negative and significant, suggesting that investors generally prefer lottery-like assets.

In sum, while our results confirm the previous findings that the anomalies related to lotteries are driven by the investor preference for lottery-like assets, both our portfolio and regression results highlight the role of CGO in lottery anomalies. In particular, the preference for lottery-like assets is much more salient when investors face prior losses but is very weak or non-existent following prior gains.

3 Inspecting the Mechanisms

In this section, we investigate several possible explanations for the pattern of lottery-related anomalies presented in the previous section. We consider the role of RDP, underreaction to past news, and general mispricing.

3.1 The Role of RDP

Most asset pricing models employ the expected utility function that is globally concave, and hence investors are uniformly risk averse. This assumption has been a basic premise of most research in finance and economics. However, building on the reference-dependent preference model by Kahneman and Tversky (1979), a large number of subsequent studies have demonstrated that RDP can better capture human behavior in decision making (e.g., Koszegi and Rabin (2006, 2007)) and can account for many asset pricing phenomena[10].

Abundant prior studies, including Odean (1998), Barber and Odean (2000, 2001, 2002), Grinblatt and Keloharju (2001), and Dhar and Zhou (2006), have documented that individual investors indeed exhibit RDP and are averse to loss realization. In addition, professional investors also exhibit RDP: see Locke and Mann (2000) 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. Moreover, RDP has psychological foundations in hedonic adaptation (see, Frederick and Loewenstein (1999)). Lastly, Rayo and Becker (2007) provide an evolutionary foundation for RDP.

Probably the most well-known RDP is the prospect theory of Kahneman and Tversky (1979), which has attracted a lot of attention in finance literature. A critical element in this theory is the reference point. The theory predicts that most individuals have an S-shaped value function that is concave in the gain domain and convex in the loss domain, both measured relative to the reference point (i.e., diminishing sensitivity). In addition, investors are loss averse in the sense that the dis-utility from losses is much higher than the utility from the same amount of gains. Thus, individuals' attitude to risk could depend on whether the outcome is below or above the reference point.[11] Finally, the mental accounting of Thaler (1980, 1985) provides a theoretical foundation for the way in which decision makers set reference points separately for each asset they own by ignoring possible interaction among these assets.

Specifically, under the assumption of the reference point being the lagged status quo, diminishing sensitivity predicts the willingness to take unfavorable risks to regain the status quo. Kyle, Ou-Yang, and Xiong (2006) shows that if the reference point is the purchase price, an investor whose investment is in losses will be risk-seeking in waiting for a price to recover before selling. A related concept, the break-even effect coined by Thaler and Johnson (1990), also suggests that following losses investors often jump at the chance to make up their losses, and the urge to break even can induce loss making investors to take gambles which they otherwise would not have taken. By breaking even, investors can avoid proving that their first judgment was wrong. As a result, assets with high skewness appear especially

[10] 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.

[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 (2013), among others.

attractive since they provide a better chance to break even, despite being in the absence of probability weighting.

However, among stocks with prior capital gains, there are two countervailing forces. On one hand, investors might still have a preference for lottery-like stocks, probably due to the standard probability weighting scheme posed by prospect theory, even despite the preference being weaker. Thus, the lottery-like stocks can still be slightly overvalued. On the other hand, the lottery-like stocks typically have higher (idiosyncratic) volatility. When facing prior gains, investors are risk-averse, even to stock-level idiosyncratic volatility due to mental accounting. Thus, the lottery-like stock can be undervalued. Overall, it is not clear which force dominates in the data. We hypothesize that these two forces tend to cancel each other, and thus, there should be no significant relationship between skewness and expected abnormal returns among stocks with prior capital gains.

In fact, using these five skewness proxies and the same brokerage data set as in Barber and Odean (2000), we show that individual investors' preference for lottery-like assets over non-lottery-assets is significantly stronger in the loss region than in the gain region.[12] Additionally, using mutual fund holding data, we find that mutual fund managers exhibit the same trading behavior. Specifically, the significant coefficients for the interaction terms between returns and skewness proxies in Tables A1 and A2 in the Appendix imply that investors exhibit a stronger preference for lottery-like assets over non-lottery assets after losses than after gains, confirming our conjecture about the role of reference points in an investor's preference for lottery-like assets.

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, and that the negative relationship between skewness and expected returns should be much more pronounced among stocks where investors have experienced losses, and thus are trying to break even. That is, the lottery-related anomalies should crucially depend on individual stocks' CGO: a negative association between expected (abnormal) returns and skewness exists among firms with a small, negative CGO, an insignificant association between expected abnormal returns and skewness exists among firms with a large, positive CGO, and more importantly, the return spread between high- and low-skewness stocks among firms with capital losses should be significantly more negative than that among firms with capital gains. This is exactly the pattern presented in section 2.

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

Lastly, prior studies also suggest other channels for these lottery-related anomalies. Probably the most prominent explanation is the overweighting of small probability events. This probability weighting effect can lead to overpricing of positively skewed assets, and thus can potentially account for anomalies related to maximum daily returns, jackpot probability and the expected skewness. Also the larger default option value of distressed firms, combined with shareholder expropriation, could account for the low return of the distressed firms since the default option is a hedge (see, e.g., Garlappi, Shu, and Yan (2008) and Garlappi and Yan (2011)).[13] In fact, the evidence from our double-sorts that the lottery anomalies are generally significant in the middle-CGO group supports the significant role of this probability weighting.

However, one key difference between our mechanism and the prior mechanisms is the heterogeneity of the lottery effect across stocks. Our mechanism implies that the lottery-related anomalies should be much more pronounced among firms with low CGO, whereas prior mechanisms typically imply 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 is homogeneous. Again, we would like to emphasize that our mechanism does not depend on probability weighting: even without probability weighting, the break-even effect and the investor's desire to avoid losses could lead to a preference for positive skewness when facing prior losses. Thus, our explanation for the lottery-related anomalies, which is based on reference points, is distinct from those based on probability weighting, which is the prevalent explanation for the preference for lottery-like assets in the existing literature (see, e.g, Barberis and Huang (2008), Bali, Cakici, and Whitelaw (2011), and Conrad, Kapadia, and Xing (2014)). Both these forces are likely to work simultaneously in investors' decision-making process, and the probability weighting would be significantly amplified by diminishing sensitivity among prior losers.

3.2 Underreaction to News

Although we have controlled for several traditional return-predictors in regression (1), it is still possible that some other forces could account for this empirical pattern. Zhang (2006), for example, argues that information may travel slowly, which can lead to underreaction to

[13] 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.

news. This underreaction effect might be stronger among firms with high arbitrage risk (or information uncertainty in Zhang's (2006) terminology). Thus, among the firms with recent good news, high arbitrage risk is likely to forecast high future returns due to the current undervaluation.

Our proxies for the lottery-like feature could be related to arbitrage risk, especially for Campbell et al's (2008) failure probability and Ohlson's (1980) bankruptcy probability.[14] Since high-CGO firms are likely to have experienced good news in the past and lottery-like firms might have high arbitrage risk, a positive relation between our proxies and return among high-CGO firms is likely to be observed in the data. On the other hand, firms with low CGO are likely to have experienced negative news and have been overpriced due to underreaction. This overpricing effect is stronger when the arbitrage risk is higher, since the underreaction effect is larger. Thus, there could be a negative relation between our proxies and return among firms with low CGO. The above argument leads to a return pattern observed in Table 2 and also implies a positive coefficient for the interaction term between CGO and skewness proxies.

To ensure that our empirical results are not mainly driven by this underreaction-to-news effect, we perform Fama-MacBeth regressions by including an interaction term between a proxy for the past news and skewness proxies. Following Zhang (2006), we use past realized returns as a proxy for news. Regression (2) in Table 6 shows that the interactions of past returns and our proxies for lottery feature are insignificant for all the skewness proxies except for the expected skewness. Thus, it is unlikely that the underreaction-to-news effect drives our key results. Indeed, after controlling for the underreaction-to-news effect, the interactions of CGO and lottery proxies remain significant with similar t-statistics. In particular, the t-statistic for the interaction between CGO and lottery proxies are 13.21 for maximum daily return, 8.01 for jackpot probability, 5.24 for expected idiosyncratic skewness, 2.21 for failure probability, and 6.02 for bankruptcy probability.

3.3 CGO as a Proxy for Mispricing

Another potential concern is that, besides being a proxy for aggregate capital gains, CGO itself is directly associated with mispricing. As documented by Grinblatt and Han (2005),

[14]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 and/or evaluate.

firms with higher CGO tend to experience higher selling pressure due to the disposition effect (characterized by 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. If our proxies for the lottery-like feature are related to arbitrage risk, then the positive relation between CGO and future returns can be amplified when firms have high skewness/default probability, leading to a positive coefficient for the interaction term between CGO and skewnesss proxies in the Fama-MacBeth regression. Thus, this is a potential explanation for the return pattern we have documented. Notice that this explanation does not rely on investors' stronger 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 mispricing measure (as compared with CGO) derived from the V-shaped disposition effect as in An (2014). 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 increase, 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 (2015) shows that stocks with large unrealized gains and losses tend to outperform stocks with modestly unrealized gains and losses. More importantly, the V-shaped Net Selling Propensity (VNSP), a more precise mispricing measure, subsumes the return predictive power of CGO.

Thus, to alleviate the concern that our CGO is a proxy for mispricing and our skewness measures are proxies for limits to arbitrage, we add an interaction term between our skewness proxies and VNSP to directly control for potential interaction effects between skewness and mispricing due to the disposition effect. Regression (3) in Table 6 shows that, $Proxy \times VNSP$ is not significant for 2 out of 5 proxies, suggesting this interaction between skewness and mispricing is empirically not a severe concern. More importantly, after controlling this effect, coefficients of $Proxy \times CGO$ remain similar in magnitude to those in regression (1), and the t-stats are positive and significant. The evidence suggests that the mispricing story does not explain the return pattern we document.

In sum, both underreaction-to-news effect and the more general mispricing story cannot account for the return pattern we have documented in Table 2. Coupled with investors' trading behavior documented in Tables A1 and A2, we believe that the strong preference for lottery-like assets after prior losses plays a role in the lottery-related anomalies.

4 Additional Robustness Checks

We now conduct a series of additional tests to assess the robustness of our results. 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 total market capitalization. Although the result based on equal-weighted regressions measures the effect of a typical firm, it might not reflect the effect of an average dollar. To alleviate this size effect, we perform the value-weighted Fama-MacBeth regressions using the same model in Table 6, where each return is weighted by the firm's market capitalization at the end of the previous month.

Second, we want to ensure that the role of reference-dependent preferences is not driven by NASDAQ stocks. Third, previous studies (e.g., Bali et al. (2005)) show that some asset pricing phenomena disappear once the most illiquid stocks are excluded from the sample. Thus, to make sure that our results are not driven by stocks with extremely low liquidity, we consider a subset of stocks that can be classified as the top 90% liquid stock according to Amihud's (2002) liquidity measure. Here, illiquidity is measured as 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 in Table 6, but now with the following alternative specifications: 1) We use weighted least square (WLS) regressions where the weight is equal to a firm's market capitalization at the end of previous month; 2) We exclude all NASDAQ stocks and use stocks listed on NYSE and AMEX only; 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 on the interaction between CGO and lottery-feature proxies are similar to those obtained in the standard Fama-MacBeth regressions in Table 6. In particular, all the t-statistics are still larger than 1.96. In addition, Table A3 in the Appendix reports the double sorting portfolio returns by excluding NASDAQ firms and illiquid firms. The role of RDP remains largely the same in portfolio returns.

In sum, the evidence in Table 7 shows that the role of the RDP in the skewness-return relation is not driven by highly illiquid stocks, NASDAQ stocks, or disproportionately affected by small firms, since both the statistical and the economic magnitude remains largely the same as before.

5 Conclusion

In this paper, we conjecture that investor's preference for lottery-like assets is state-dependent. In particular, the break-even effect and loss aversion suggest that following losses investors often jump at the chance to make up their losses, and the urge to break even can induce investors who face prior losses to take gambles which they otherwise would not have taken. Thus, assets with high skewness appear especially attractive since they provide a better chance to break even. As a result, with mental accounting, investors' preference for lottery-like assets is much stronger among stocks where average investors are in losses than among stocks where average investors are in gains.

More importantly, we show that due to changes in the preferences for lottery-like assets, 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 prior capital gains in these investments, the underperformance of lottery-like assets is either extremely weak or even reversed. Thus, we provide a unified framework to understand many lottery-related anomalies in the literature, such as those related to maximum daily returns, jackpot probability, expected skewness, failure probability, and bankruptcy probability.

References

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

An, Li, 2015, Asset Pricing When Traders Sell Extreme Winners and Losers, *Review of Financial Studies*, forthcoming.

Anderson, Anne-Marie, Edward A. Dyl, 2005, Market Structure and Trading Volume, *Journal of Financial Research* 28, 115–131.

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

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, 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, Ming Huang, and Richard H. Thaler, 2006, Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing, *American Economic Review*, 96, 1069–90.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2015, Prospect Theory and Stock Returns: An Empirical Test, *Working Paper*, Yale University and HKUST.

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, 2014, Anchoring on Credit Spreads, *Journal of Finance*, forthcoming.

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, and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

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.

Friedman, Milton, and L.J. Savage, 1948, The utility analysis of choices involving risk, *Journal of Political Economy* 56, 279–304.

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.

Genesove, David, and Christopher Mayer, 2001, Loss Aversion and Seller Behavior: Evidence from the Housing Market, *Quarterly Journal of Economics* 116, 1233-1260.

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

Gomes, Francisco, 2005, Portfolio Choice and Trading Volume with Loss-Averse Investors, *Journal of Business* 78, 675–706.

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.

Koszegi, Botond, and Matthew Rabin, 2006, A Model of Reference-Dependent Preferences, *Quarterly Journal of Economics* 121, 1133–1165.

Koszegi, Botond, and Matthew Rabin, 2007, Reference-Dependent Risk Attitudes, *American Economic Review* 97, 1047–1073.

Kumar, Alok, 2009, Who Gambles in the Stock Market? *Journal of Finance* 64, 1889–1933.

Kyle, A. S., H. Ou-Yang, and W. Xiong, 2006, Prospect theory and liquidation decisions, *Journal of Economic Theory* 129, 273–288.

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.

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..

Zhang, Yijie, 2005, Individual Skewness and the Cross-section of Average Returns, Unpublished Working Paper, Yale University. 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, $Ret_{-12,-2}$ is the cumulative return over the past year with one month gap, $Ret_{-36,-13}$ 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}	$Ret_{-12,-2}$	$Ret_{-36,-13}$	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)	(11.7)	(-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)	(-3.7)	(16.55)	(-1.64)	(-5)	(11.02)	(0.42)	(-2.73)	(14.8)

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)	(2.3)	(7.5)	(-4.15)	(-2.16)	(2.3)	(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)	(4.7)				
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)	(5.9)				

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	(-4.7)	(-2.23)	(0.83)	(5.99)	(-3.45)	(-2.23)	(0.29)	(4.09)	(-1.88)	(-1.75)	(-0.7)	(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)	(-0.9)	(-0.94)	(3.1)	(-2.94)	(0.85)	(-0.4)	(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)	(-0.3)	(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, 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). 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 December 2014 for Skewexp. In the cases of CGO^{FR} and Residual CGO^{FR}, the sample period is from January 1980 to October 2014 for Maxret, Oscorep, Jackpotp and Deathp, and from January 1988 to October 2014 for Skewexp. The t-statistics are calculated based on the heteroskedasticity-consistent standard errors of White (1980).

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)	(-0.3)	(4.36)	(-6.12)	(-0.64)	(3.64)	(-4.7)	(-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)	(2.3)	(-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)	(5.9)	(-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).

	(I) Equal-Weighted			(II) Lag-Ret-Weighted			(III) Conditional Sort		
Proxy	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1
Maxret	-1.81	0.08	1.88	-1.88	0.09	1.97	-1.74	0.25	1.99
	(-13.86)	(0.57)	(10.78)	(-14.7)	(0.64)	(11.2)	(-8)	(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)	(-8.1)	(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)	(4.7)	(-4.07)	(1.36)	(5.21)	(-4.24)	(-1.13)	(2.7)
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)	(-7.1)	(-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, $Ret_{-12,-2}$ is the cumulative return over the past year with a one-month gap, $Ret_{-36,-13}$ 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 Conrad et al.(2014), Skewexp is the expected idiosyncratic skewness 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$\times$ 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$\times Ret_{-12,-2}$			-0.054		-0.054		0.084		0.068
			(-2.33)		(-2.15)		(0.72)		(0.57)
Proxy$\times$ VNSP				0.160	0.228			1.257	1.238
				(1.93)	(2.55)			(2.81)	(2.7)
Ret_{-1}	-0.062	-0.056	-0.056	-0.059	-0.059	-0.051	-0.051	-0.055	-0.055
	(-15.8)	(-13.45)	(-13.54)	(-14.16)	(-14.24)	(-12.28)	(-12.39)	(-12.98)	(-13.06)
$Ret_{-12,-2}$	0.008	0.008	0.011	0.006	0.009	0.008	0.007	0.006	0.005
	(5.4)	(5.4)	(6.4)	(4.06)	(5.16)	(4.93)	(3.21)	(3.78)	(2.39)
$Ret_{-36,-13}$	-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)	(-2.2)	(-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)	(2)	(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)	(-1.6)	(-1.4)	(-1.35)	(-1.15)	(-1.12)

	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)	(-0.28)	(-1.41)	(-1.44)	(0.92)	(1)	(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)	(-9.3)	(-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×$Ret_{-12,-2}$		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)
$Ret_{-12,-2}$	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)	(5.4)	(5.05)	(4.11)	(3.85)
$Ret_{-36,-13}$	-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)	(-2.6)	(-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
	(-1.1)	(-1.19)	(-1.15)	(-1.19)	(-2.8)	(-2.8)	(-2.79)	(-2.7)	(-3.27)	(-3.21)	(-3.36)	(-3.3)
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)	(0.1)	(0.1)	(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)	(-1.9)

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).

	(I) WLS					(II) Excluding NASDAQ stocks					(III) Excluding top illiquid decile				
Proxy =	Maxret	Jackpotp	Skewexp	Deathp	Oscorep	Maxret	Jackpotp	Skewexp	Deathp	Oscorep	Maxret	Jackpotp	Skewexp	Deathp	Oscorep
CGO	-0.013	-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
	(-1.3)	(-5.17)	(-0.82)	(-1.94)	(-2.91)	(-2.2)	(-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)	(2.6)	(3.3)	(8.03)	(3.51)	(2.9)	(2.46)	(4.72)	(8.68)	(4.44)	(3.75)	(3.67)	(3.49)
Proxy× $Ret_{-12,-2}$	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
	(2.1)	(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)	(-6.8)	(-3.2)	(-8.38)	(-9.22)	(-10.97)	(-10.2)	(-3.97)	(-12.4)	(-13.16)	(-13)	(-12.13)	(-6.63)	(-14.48)	(-14.5)
$Ret_{-12,-2}$	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)	(2.6)	(2.73)	(3.72)	(3.53)	(1.63)	(1.39)	(0.85)	(3.36)	(4.48)	(2.63)	(1.65)	(2.39)	(3.76)
$Ret_{-36,-13}$	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)	(-1.1)	(-0.9)	(-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)	(1.4)	(-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)	(0.85)	(0.2)	(-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
	(-0.5)	(-0.31)	(0.21)	(-0.9)	(-1.2)	(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)

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 Grinblatt 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× CGO	0.282	0.583	0.010	2.961	0.110
	(9.19)	(3.01)	(2.97)	(1.01)	(2.82)
Proxy×$Ret_{-12,-2}$	-0.053	0.088	0.006	1.075	0.008
	(-2.12)	(0.73)	(2.9)	(0.41)	(0.19)
Proxy× VNSP	0.225	1.226	0.001	2.312	0.267
	(2.53)	(2.68)	(0.12)	(0.26)	(2.05)
RetVol× 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)
$Ret_{-12,-2}$	0.009	0.004	0.001	0.003	0.005
	(4.98)	(2.16)	(0.39)	(1.63)	(3.32)
$Ret_{-36,-13}$	-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)	(-1.2)	(-2.7)	(-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).

	(I) Excluding NASDAQ Stocks			(II) Excluding Top Illiquid Decile		
Proxy	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)	(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)	(2.7)
Deathp	-1.35	-0.12	1.23	-1.98	-0.44	1.54
	(-5.13)	(-0.5)	(3.97)	(-7.1)	(-2.3)	(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).

	Top 25% IO			Bottom 25% IO			Top - Bottom IO		
Lottery Proxy	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1
Maxret	-0.94	0.21	1.15	-2.43	-0.22	2.21	1.49	0.43	-1.06
	(-3.48)	(0.67)	(3.65)	(-8.58)	(-0.68)	(5.15)	(4.23)	(1.15)	(-2)
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)	(4.3)	(2.59)	(-0.45)	(-2.68)
	Top 25% Price			Bottom 25% Price			Top - Bottom Price		
Lottery Proxy	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1	CGO1	CGO5	C5-C1
Maxret	-0.46	0.62	1.08	-3.00	-1.11	1.88	2.54	1.73	-0.81
	(-2.3)	(3.07)	(4.61)	(-9.5)	(-4.13)	(5)	(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)	(2.6)	(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)	(2.9)	(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)	(-2.1)

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).

	High Sentiment Periods			Low Sentiment Periods			High-Low Sentiment Periods		
Lottery Proxy	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)	(0.2)	(2.44)
Oscorep	-1.59	0.36	1.95	-0.97	0.00	0.97	-0.63	0.36	0.99
	(-3.95)	(1)	(3.68)	(-1.77)	(0)	(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 with having 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 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 starts from April 1980 to October 2014.

Jackpotp: The predicted jackpot probability is constructed from the baseline model in Conrad et al. (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 et al.(2009) (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 et al.(2009)'s 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 et al.(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:

$$\begin{aligned} Distress_t = & -9.16 - 20.26 NIMTAAVG_t + 1.42 TLMTA_t - 7.13 EXRETAVG_t \\ & + 1.41 SIGMA_t - 0.045 RSIZE - 2.13 CASHMTA_t + 0.075 MB_t - 0.058 PRICE_t, \end{aligned}$$

where

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

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 :

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

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

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