Review of Finance (2012) 16: 685–731
doi: 10.1093/rof/rfr006
Advance Access publication: May 19, 2011

Gambling Preference and the New Year Effect of Assets with Lottery Features*

JAMES S. DORAN, DANLING JIANG†, and DAVID R. PETERSON
College of Business, The Florida State University

Abstract. This paper shows that a New Year's gambling preference of individual investors impacts prices and returns of assets with lottery features. January call options, especially the out-of-the-money calls, have higher retail demand and are the most expensive and actively traded. Lottery-type stocks outperform their counterparts in January but tend to underperform in other months. Retail sentiment is more bullish in lottery-type stocks in January than in other months. Furthermore, lottery-type Chinese stocks outperform in the Chinese New Year's Month but not in January. This New Year effect provides new insights into the broad phenomena related to the January effect.

JEL Classification: G12, G14

1. Introduction

Individuals' preference to gamble is often an explanation for a number of aspects of individual financial decision making, such as the purchase of both insurance and lotteries (Friedman and Savage, 1948; Markowitz, 1952), portfolio underdiversification (Statman, 2004), and portfolio overweighting on lottery-like securities (Kumar, 2009). Recent theoretical development (Shefrin and Statman, 2000; Brunnermeier, Gollier, and Parker, 2007; Mitton and Vorkink, 2007; Barberis and Huang, 2008) advances this notion into asset pricing. These theories show that

* We appreciate helpful comments and suggestions on earlier drafts from Hal Arkes, Don Autore, Gurdip Bakshi, Robert Battalio, Gary Benesh, Jie Cao, Honghui Chen, Mike Cliff, Gonul Colak, John Griffin, David Hirshleifer, Irena Hutton, Narasimhan Jegadeesh, Alok Kumar, Chris James, Sonya Lim, Xiaoding Liu, Roger Loh, Mahendrarajah Nimalendran (Nimal), Jay Ritter, Hersh Shefrin, K.R. Subramanyam, Rossen Valkanov, Yexiao Xu and seminar participants at The Florida State University and University of Florida, conference participants at the 2008 BDRM conference hosted by Rady School of Management in UCSD, 2008 China International conference in finance in Dalian, China, and 2008 FMA at Dallas, Texas. We appreciate Ji-Woong Chung and Kuan-Hui Lee for kindly answering questions about Datastream, and Ansley Chua, Jordana Gordon, Lanis Heuiser, Dan Lawson, and Maurizio Mendoza for helpful research assistance. This paper subsumes the earlier version entitled "Gambling in the New Year? The January Idiosyncratic Volatility Puzzle." All errors remain ours.
†All correspondence to be addressed to Danling Jiang

a gambling preference by some market participants can cause overpricing of securities with lottery features.

Evidence from gambling, risk taking, and portfolio rebalancing further suggests that the gambling preference of individuals may have a stronger price impact at the turn of the year.[1] First, individuals may exhibit stronger gambling mentality in the New Year. Anecdotal evidence shows that individuals around the world actively engage in lottery plays, casino gambling, race betting, and home gaming as a way to celebrate the New Year. In the USA, revenues in the Las Vegas strip and interstate lottery sales on Mega Millions and Powerball are unusually high at the turn of the New Year. Second, experimental work on individual decision making finds that people engage in risk-seeking activities after observing payoff outcomes in prior gambling rounds (Thaler and Johnson, 1990) and that multiperiod financial decisions are commonly evaluated in intertemporal mental accounts (Thaler, 1985), which can be labeled by year. As a natural starting point for a new round of gambling/investing, risk taking likely strengthens in the New Year. Third, since at the turn of the New Year investors usually receive annual reports from mutual funds, prepare taxes, and receive bonuses, they most likely evaluate and rebalance portfolios, make New Year's resolutions, and choose new investments (Benartzi and Thaler, 1995).[2] Therefore, at the turn of the year, investors' preference to gamble is most likely to influence asset prices through their unusually strong buying activities.

In this paper, we study whether the New Year's gambling mentality of individual investors impacts prices, returns, and trading volume of financial assets with lottery features. We hypothesize that investors are most likely to place lottery-type bets in financial markets at the start of a New Year, elevating prices and returns of lottery-like options and stocks at that time. Furthermore, in markets such as China, lottery-like stocks should outperform at the start of the Chinese New Year but not necessarily in January. We find strong evidence supporting the above hypotheses. The pricing impacts are economically and statistically significant. We also show that, in the USA, such pricing impacts correspond to higher demand for lottery-type options and equity by retail investors in January. Our results provide novel evidence that gambling preference has aggregate price impacts and offer fresh insights into several well-known January effects.

[1] See Section 2 for more detailed review of the evidence.

[2] Benartzi and Thaler (1995) suggest that receiving annual reports and filing taxes at the year-end likely force investors to evaluate and reallocate their portfolios. There is also evidence that individual investors tend to sell stocks toward the year-end and purchase early in January to rebalance their portfolios (Ritter 1988; Sias and Starks 1997; D'Mello, Ferris, and Hwang 2003; Starks, Yong, and Zheng 2006).

We begin our tests with the US option markets. Behavioral theories (Shefrin and Statman, 2000; Barberis and Huang, 2008) suggest that out-of-the-money (OTM) call options are natural candidates for gambling purposes; they are cheap and have highly skewed payoffs. If, at the turn of the year, investors desire to purchase lottery-type assets, they will overdemand OTM calls and drive up the prices and volume of these securities. Using at-the-money (ATM) calls on the same stocks as benchmarks, we show that the implied volatility (measuring the relative expensiveness of options) and volume on the OTM calls are significantly greater in January than in other months, revealing novel seasonality consistent with the New Year's gambling mentality of investors.

We also examine trading behavior from accounts of clearing member firms (which mostly include big institutions) versus customers (including, among others, retail investors) in option markets. We expect that, relative to firms, customers more likely place open buy orders on call than on put options, particularly in January.[3] This is consistent with the notion that unsophisticated investors usually bet on calls, rather than puts, for upside potential (Shefrin and Statman, 2000; Statman, 2002). We find that customers, particularly those who place small bets per trade, open significantly more buy contracts on calls than puts and particularly so in January. Such differentials in preference between January and non-January months are absent for open put buys. In other words, our evidence suggests that in the option markets, small retail investors exhibit strong gambling preferences in the New Year and reveal such preferences through buying OTM calls.

Does the New Year's gambling mentality in the option markets translate to equity markets and impact stock prices and returns? We explore this question in USA and China stock markets. In the USA, we employ three measures of stock lottery features following Kumar (2009) and Boyer, Mitton, and Vorkink (2010): low stock price, high idiosyncratic return volatility, and high expected idiosyncratic return skewness.[4] In addition to using each independently, we also form a composite lottery-feature index incorporating all three features. We show that stocks with strong lottery features significantly outperform their counterparts in January; in other months, such outperformance is attenuated and commonly reversed.

Over the 44 January months from 1964 through 2007, an equal-weighted (value-weighted) hedge portfolio that is long the highest lottery-feature index quintile and short the lowest index quintile generates a mean January return of 11.48% (7.21%)

[3] Open buys refer to new purchases of options and do not include buying to cover previous short option positions.

[4] Kumar (2009) suggests that investors begin searching for lottery-type stocks among those with low prices and then examine those with high skewness. High return skewness indicates a small probability of winning a large "jackpot," while high return volatility likely inflates the perception of the likelihood that the extremely large payoff is paid.

and a monthly Fama–French three-factor alpha of 8.51% (4.13%). The equal-weighted portfolio has only one negative January return. The January outperformance of lottery-type stocks is robust to adjustments for the bid–ask spread (Keim, 1989), delisting bias (Shumway, 1997), elimination of stocks trading below $5, and controls for the January effects in firm size, book-to-market equity, past short- and long-run returns, and loadings on a set of common factors. The New Year effect is concentrated within a 20-day window surrounding New Year's Day and cannot be fully explained by tax-loss selling or institutional window-dressing and risk shifting.

Furthermore, using a retail sentiment measure developed by Barber, Odean, and Zhu (2009) based on the Institute for the Study of Securities Market/Trade and Quote (ISSM/TAQ) data, we show that only in (early) January are retail traders significantly more bullish on lottery-type stocks than stocks with opposite characteristics. Retail traders tend to rebalance their portfolios by selling all kinds of holdings toward the year-end and purchasing lottery-type stocks early in January. The January price run-ups of lottery-type stocks are significantly greater among those with stronger bullish retail sentiment or those more actively traded by retail investors. This contrasts with institutional trades that exhibit no January seasonality. Additionally, lottery-type stocks on which retail investors are bullish in January deliver negative future returns for the remainder of the year, while those favored by institutional investors do not. This evidence is consistent with the finding by Kumar (2009) that retail investors exhibit stronger gambling preferences than institutional investors in portfolio allocation.

In the China stock markets, we examine if gambling mentality impacts stock returns in the New Year but not necessarily in January unless the two coincide. This is important for distinguishing the gambling preference–based hypothesis from traditional hypotheses for broad January effects that involve tax-loss selling (Starks, Yong, and Zheng, 2006), institutional window dressing (Haugen and Lakonishok, 1992), and institutional risk shifting (Ng and Wang, 2004). Chinese stock markets provide an ideal setting to test these competing hypotheses. Chinese celebrate the traditional Chinese New Year's Day based on the lunar calendar, also known as the "Spring Festival," more seriously than January 1.[5] Chinese traditionally gamble in the New Year. There is no income or capital gains tax on stock trading and almost all investors are retail.[6] Both the tax-loss selling and institutional

5 The Chinese Spring Festival is a single day and equivalent to the US New Year's Day. It is based on the lunar calendar and usually occurs somewhere from mid-January to mid-February, depending on the year.

6 Chinese stock investors only pay a stamp tax for each stock transaction. In 1998 mutual funds held only 2% of the Chinese tradable A-shares and in March of 2006, their holdings rose to 14.4%. This explosive growth, however, occurs only after 2003 when the first national legislation on securities investment funds, the Law on Securities Investment Funds, went into effect (Xi (2006)).

trading–based hypotheses predict no January or New Year effect. In contrast, the gambling-preference-based hypothesis predicts that the Chinese market as a whole and the lottery-type Chinese stocks outperform at the start of the Chinese New Year but not necessarily in January. We find strong evidence for these two predictions.

Our hypotheses and findings provide unique insights into a set of long-standing phenomena related to the January effect in stock/bond returns (Rozeff and Kinney, 1976). The predominate explanation for the January effect is tax-loss selling. But, the January effect is found in countries or time periods with no capital gains taxes (Kato and Schallheim, 1985; Van den Bergh and Wessels, 1985) and in countries with tax years ending in a month other than December (Gultekin and Gultekin, 1983; Brown et al., 1983). The January effect occurs for noninvestment-grade bonds but not for investment-grade bonds (Maxwell, 1998). These findings fit into our interpretation based on investor's gambling preference in the New Year. In the USA, adding lottery feature variables (particularly skewness) in Fama–MacBeth regressions reduces the magnitude of, or even reverses, the sign of the coefficients on firm size.

Our results are not, however, a repackaging of existing January effects. Our lottery measures show predictive power after controlling for January-related firm characteristics. Lottery-type stocks do well in January from the early eighties through late nineties, a period during which the small-firm-in-January effect dissipated (Schwert, 2003; Haug and Hirschey, 2006). More importantly, our findings about January effects of idiosyncratic volatility and skewness, of OTM versus ATM calls, the retail sentiment revealed in the order imbalance, and the New Year effect of Chinese stocks are all novel. While the gambling preference–based hypothesis provides a coherent story for our results, the findings are interesting irrespective of the interpretation.

The remainder of this paper is organized as follows. Section 2 summarizes the motivational literature. Section 3 describes the data. Section 4 presents the empirical results from US option and stock markets and China stock markets. Section 5 discusses the implications for the wealth transfer implied by our results and the survival of gamblers. Section 6 summarizes and concludes.

2. Motivation and Hypotheses

2.1 GAMBLING PREFERENCE AND ASSET PRICES

The notion that individuals have a preference to speculate with part of their wealth emerged over 50 years ago. Both Friedman and Savage (1948) and Markowitz (1952) note that it is puzzling that individuals often buy insurance as well as lotteries. They question why individuals exhibit both risk-aversion and risk-seeking behavior. Behavioral portfolio theory developed by Shefrin and Statman (2000) suggests one possibility. Investors view financial assets as pyramids, purchasing insurance for downside protection, diversified mutual funds to ensure their current social rank, and securities with speculative features for an upside potential. Based on the level of investor aspirations to move upward in social class, they can choose aggressive individual stocks, call options, or lotteries. Therefore, gambling preferences are not caused by risk seeking but by aspiration.

Alternative theoretical work suggests other reasons for the preference to gamble. For instance, Barberis and Huang (2008) suggest that investors are willing to pay for skewness because they overweight events that have extremely small probabilities, which is an aspect of prospect theory (Kahneman and Tversky, 1979). Brunnermeier, Gollier, and Parker (2007) model the preference for skewness as an outcome of investors being overly optimistic about the probability of good states. In the model of Mitton and Vorkink (2007), investors have heterogeneous preferences for skewness. The above models all conclude that the preference for skewness impacts equilibrium prices; securities with speculative features are overpriced and have lower expected returns. These models help explain stock market phenomena such as the pricing of OTM options (Bollen and Whaley, 2004) and the underperformance of high idiosyncratic volatility securities (Ang *et al.*, 2006).

Kumar (2009) shows that lottery-type stocks are overweighted in portfolios of retail investors but not in those of institutional investors. Zhang (2005) and Boyer, Mitton, and Vorkink (2010) show that measures of expected skewness of idiosyncratic stock returns are negatively related to future stock returns, supporting the prediction that highly skewed stocks are overpriced due to investor preferences for skewness. Our paper is most related to the above research. We find that high idiosyncratic skewness stocks outperform in January, but underperform in other months, suggesting that overpricing of highly skewed stocks largely occurs in January and is corrected in the remaining months of the year. We also go beyond the above empirical literature by studying the gambling preference–induced seasonality in option markets and Chinese stock markets.

2.2 GAMBLING PREFERENCE AT THE TURN-OF-THE-NEW-YEAR

There are several reasons for why we expect investors to engage in excess buying in lottery-type securities at the turn of the year. First, investors receive bonuses over that window and part of the new money can flow into the financial market.[7] If an

[7] Employees on Wall Street typically get their bonus numbers in the first two weeks of December–with the cash coming early in the New Year. See "Goldman Chiefs Give Up Bonuses," by Susanne Craig, *Wall Street Journal*, 11/17/2008. Year-end bonuses and early contributions for the New Year are considered as factors contributing to the January effect of stock market. See "Investors Advised To Beware Gold Fever," by Russ Wiles, *Chicago Sun-Times*, March 17, 2003.

average investor has a preference to gamble, the new money will flow disproportionally more into lottery-type securities, thus elevating the prices of these securities. Second, existing literature shows that investors, particularly individual investors, tend to rebalance their portfolios at the turn of the year (Ritter, 1988; Sias and Starks, 1997; D'Mello, Ferris, and Hwang, 2003; Starks, Yong, and Zheng, 2006). Investors tend to sell stocks toward the year-end, partly motivated by tax-loss selling, and purchase stocks early in the New Year.[8] Such portfolio rebalancing is strengthened by the fact that investors usually receive annual reports on their retirement account and mutual fund investments toward the year-end. Thus, it is natural to evaluate portfolios and rebalance them accordingly at the turn of the year.

Third, experimental evidence from psychology and financial decision making suggests that investors may exhibit different risk-seeking behavior at the turn of the New Year. For instance, individuals tend to change their risk-taking tendency when decisions are framed in a multiperiod setting. Thaler and Johnson (1990) show that after experiencing prior gains in a lottery play, individuals become more risk seeking in subsequent plays; this is termed the "house money effect." After experiencing losses but being offered a chance to break even, individuals are also more willing to gamble; this is termed the "break-even effect." Evidence consistent with these effects is found for various market participants by Ackert *et al.* (2006), Coval and Shumway (2005), O'Connell and Teo (2007), and Liu *et al.* (2010).[9] The path-dependent risk-taking behavior is also shown to be crucial in understanding a set of stylized facts in financial markets, including the equity premium puzzle, excess volatility, and aggregate return predictability (Barberis, Huang, and Santos, 2001). One explanation is that multiple-period financial decisions are evaluated and labeled in intertemporal mental accounts (Thaler, 1985). Depending on how investors integrate and segregate prior gains and losses across multiple mental accounts, they exhibit differing risk-taking behavior in subsequent trading (Weber and Camerer, 1998; Arkes *et al.*, 2008). Mental accounting by investors is important for understanding individual investor trading (Kumar and Lim, 2008) and stock return anomalies such as the value effect (Barberis and Huang, 2001). If the turn-of-the-year serves as the starting point for the new round of gambling/investing, which is coupled with annual bonuses working as the house money, then we expect more frequent increases in risk taking at the start of the New Year.

[8] However, tax-loss selling itself is not sufficient to explain why investors park their selling proceeds until after the New Year to invest in stock markets Ritter (1988).

[9] The finding of Coval and Shumway (2005) that morning losses of CBOT bond traders encourages their afternoon risk-taking is consistent with the break-even effect; prior losses encourage more subsequent risk-taking when there is a chance to break even. Under their setting, this behavior is partly induced by the incentive mechanism that these traders' performance is summarized and presented on a daily basis, and therefore they tend to take on extra risks to avoid losses by the end of the day.

Fourth, anecdotal evidence suggests that people actively gamble at the New Year. Section A of the online Supplementary Appendix shows an incomplete list of such anecdotal evidence from 1965 to 2008, collected from several lottery-related Web sites, Lexis-Nexis Academic Universe (available mainly from 1980), and supplemented by that from The New York Times for the pre-1980 period.[10] The excerpts are listed for the USA and several ethnicities and countries, including Chinese, Indian, Jewish, Australia, Canada, Greece, Japan, New Zealand, Spain, and Turkey.

The anecdotal evidence collectively shows that, first, stock market investors often rebalance their portfolios at the turn of the year, with a tendency to invest some money in securities with speculative features. This behavior is linked to the appreciation of over-the-counter and low-price stocks at the beginning of January. Second, in the USA and other countries, people elevate casino gambling, bet on races, and play lotteries as a way of celebrating Christmas and the New Year. For instance, for casinos in Las Vegas and the Atlantic City, New Year's Eve has been the biggest night of the year. For casinos and Internet gaming, the period surrounding the Lunar (Chinese) New Year is becoming lucrative due to the gambling preference of Chinese. The New Year's Eve (Millionaire) lottery raffles, introduced in recent years, have gained notable popularity in the USA. Race betting, casino gaming, and lottery sales appear unusually popular from Christmas to the New Year in Australia, Canada, New Zealand, and Spain. Gambling around the New Year has been a tradition to test good luck for the entire year for Chinese, Indians, Greeks, and Turks. Taken together, the anecdotal evidence suggests that the gambling mentality, either through gaming or investing, appears strong around the New Year.

2.3 SEASONALITY IN LAS VEGAS GAMING AND MEGA-MILLION/POWERBALL LOTTERY PLAYING

Since the evidence suggesting that investors are likely to gamble more in the New Year is largely anecdotal, we test this more formally using three sets of direct gaming data: the gaming revenue from the Las Vegas Strip, which is the largest location of US casinos in terms of annual gaming revenue, and the lottery sales on Mega Millions and Powerball, which are the two earliest and most popular interstate lotteries in the USA.[11] Our purpose is to examine whether casino gaming or lottery plays are more active in January than other months.

[10] The Supplementary Appendix is available as "supplemental material" on the journal's homepage www.revfin.org and from the publisher web site http://rof.oxfordjournals.org.

[11] Mega Millions began on August 31, 1996 and was initially known as the Big Game. The first drawing took place on September 6, 1996. Powerball is run by the Multi-State Lottery Association, formed on September 16, 1987. On November 2, 1997, there was a change in the Powerball payoff, which is the start date of the data we use for Powerball.

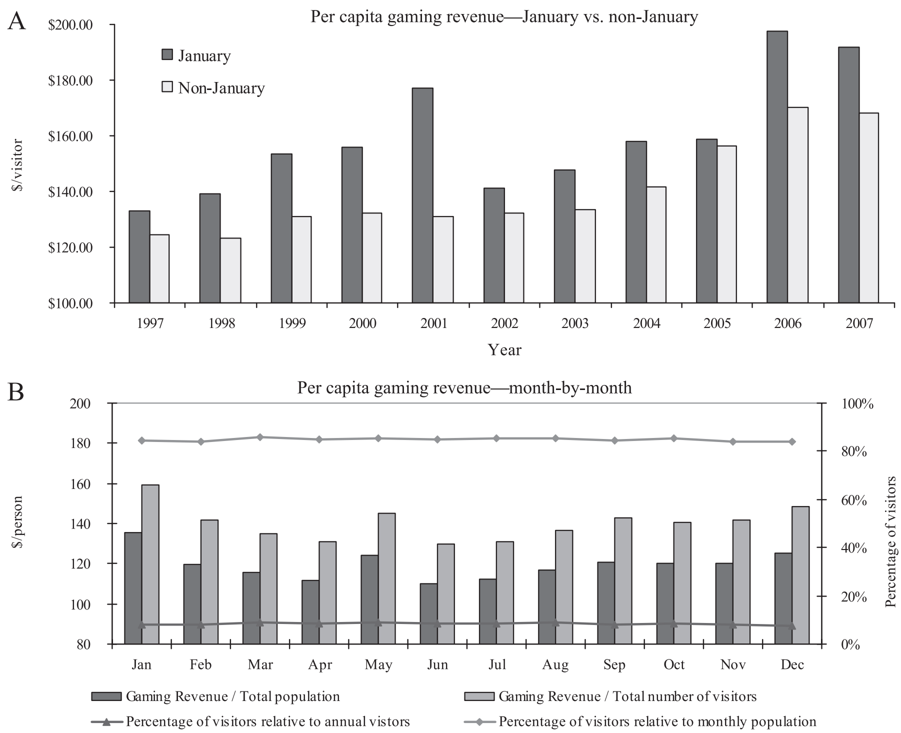

Figure 1. January seasonality of the per capita gaming revenue from the Las Vegas Strip. This figure shows that casino gambling in Las Vegas exhibits January seasonality. Panel A depicts the year-by-year average monthly per capita gaming revenue from the Las Vegas Strip in January and non-January months over the period 1997–2007. The per capita gaming revenue is defined as the monthly gaming revenue over the total number of monthly visitors in Las Vegas. Panel B depicts the month-by-month per capital gaming revenue, averaged across the eleven-year period, and the percentage of monthly visitors relative to annual visitors and to total monthly population. The per capita gaming revenues in Panel B are defined as monthly gaming revenues over either total monthly population (including local residents and visitors) or the total number of monthly visitors.

The Las Vegas Strip data is from the State of Nevada Gaming Control Board (www.gaming.nv.gov) from 1996 to 2007. The visitor statistics come from the Las Vegas Convention and Visitors Authority (www.lvcva.com), and the number of local residents is from the US Census Bureau. Panel A of Figure 1 plots the average dollar gamed per visitor on the Strip for January and other months from 1997 to 2007. The per capita gaming revenue is the total gaming revenue from the Strip divided by the number of visitors, which adjusts for visitation seasonality. The per visitor gaming revenue is higher in January than other months for all years. On average, a visitor gambles \$159.50 in January and \$140.32 per month in all other months with an average difference of \$19.18 ($t = 5.42$) or 14% more in January.[12]

Panel B further shows that the gaming activity is highest in January, regardless of whether the per capita gaming revenue is adjusted by the total population or total number of visitors. The high gaming activity in January is not caused by a larger number of gamblers in this month. In fact, there is little variation in the number of visitors to Las Vegas each month.[13] Therefore, the evidence suggests a stronger gambling mentality around the New Year.

To study whether lottery players in the USA show a New Year gambling sentiment, we study whether the lottery sales change in Mega Millions and Powerball is unusually large in early January, controlling for other factors such as the size of the jackpot, the number of states participating, and the event that a jackpot was won. We obtain from lottery.com the nationwide total sales data, available from September 1996 through May 2010 for Mega Millions and from November 1997 through May 2010 for Powerball. Both lottery games are drawn biweekly, every Tuesday and Friday for Mega Millions and Wednesday and Saturday for Powerball, and the sales are reported on the drawing dates. As of 2010, there are 42 states playing Mega Millions and 44 playing Powerball. Over our sample period, the sales follow an upward trend. To perform regression analyses on a stationary time series, we compute and use as the dependent variable the change in sales between two consecutive drawing dates. In addition to studying the nationwide total lottery sales, we also examine the sales change within individual participating states in Mega Millions, with data available from January 2003 through May 2010.[14]

In Table I, we report the ordinary least square regression results of the lottery sales change on the New Year dummy, which takes the value of 1 for drawing dates between January 1and January 20 and 0 otherwise.[15] We control for the change in jackpot size during the two preceding reporting dates, the number of states participating during the reporting period, and, in the second specification, a December dummy, which takes a value of 1 if the drawing date falls in December. Both

[12] Since the number of visitors is fairly stable throughout the year, the total gaming revenue is also the highest in January. In other words, at both the individual and the aggregate level, gambling mentality appears to be strongest at the-turn-of-the-year.

[13] The percentage of visitors in January is neither the highest nor the lowest among the months and visitors consist of a fairly stable fraction of the total population across all months due to the "smoothing out" efforts of Las Vegas local private and public sectors (Cargill and Eadington 1979).

[14] The individual-state Mega Millions sales data are available for thirteen participating states: the original six participating states in 2003, Georgia, Illinois, Maryland, Massachusetts, Michigan, Virginia, together with New Jersey, New York, Ohio, Texas, Virginia, and Washington joining during the sample period.

[15] We obtain similar results if the New Year dummy is defined including sales reported in the entire January, or reported in the last week in December through January.

Table I. Regression of lottery sales

This table reports the results of regressions of the lottery sales change (in millions) between two draw dates on the New Year dummy and other control variables. The dependent variable in Regressions (1) and (2) is the total sales change in Mega Millions, in Regression (3), it is the individual state sales change in Mega Millions, and in Regressions (4) and (5), it is the total sales changes in Powerball. The sales change is the difference in sales reported in two consecutive draw dates. The draw dates are every Tuesday and Friday for Mega Millions and Wednesday and Saturday for Powerball. Among the independent variables, the New Year dummy takes the value of 1 for a draw date between January 1 and January 20 and 0 otherwise, the change in jackpot size is measured during the two preceding draw dates, the number of states refers to the number of states participating in the sales, and the December dummy is set to one if the draw date falls into December. Regression (3) uses a panel data with thirteen states that release the state sales data by controlling for the state fixed effects from January 2003 through May 2010. The data for the Mega millions lottery cover the period from September 1996 through May 2010. The data for the Power balls lottery cover the period from November 1997 through May 2010. Individual state sales data are available starting in January 2003. The sample excludes the observation on the next draw date immediately following the draw date when the lottery is won. Robust *t* statistics are reported in parentheses.

Lottery type sample period regression	Mega Millions (1996–2010)		Individual state Mega Millions (2003–2010)	Power Ball (1997–2010)	
	(1)	(2)	(3)	(4)	(5)
New Year dummy	1.024 (3.79)	1.008 (3.75)	0.019 (2.11)	0.913 (2.02)	0.928 (2.06)
Change in jackpot size	0.632 (14.94)	0.632 (14.94)	0.055 (18.05)	0.768 (10.93)	0.767 (10.93)
Number of states	−0.382 (−10.31)	−0.382 (−10.30)		−0.643 (−7.85)	−0.642 (−7.83)
December dummy		0.110 (0.43)			0.321 (0.66)
Intercept	761.934 (10.30)	761.684 (10.30)	−0.413 (−11.89)	1,283.29 (7.84)	1,282.28 (7.83)
State fixed effect	No	No	Yes	No	No
Number of observations	1,203	1,203	7,093	1,176	1,176
R^2 (%)	82	82	56	71	71

specifications are used for the nationwide sales change in Mega Millions and Powerball, respectively. When the individual-state lottery sales change is used as the dependent variable, however, we add the state fixed effects and exclude the number of states participating. Finally, we exclude the sales change occurring immediately following a winning draw because sales are typically at their lowest and the sales change is highly negative during that reporting period.[16]

[16] The size of the lottery payout is tied to the sales of tickets, where approximately 30-32% of the sales are paid out. Thus, a winning draw usually has a large negative effect on the immediately subsequent sales.

The results in Table I depict a clear picture. For both Mega Millions and Powerball, there is a significant positive relation between the sales change and the New Year dummy. The coefficient on the New Year dummy suggests an approximately \$1 million per half-week (or over \$6 million in the first 3 weeks of January) abnormal increase in nationwide sales. Not surprisingly, the change in the size of jackpot in the prior period is always positive and significant, while the number of states participating is negative since newly added states tend to have fewer sales per period than the original states. The coefficient on the December dummy is always insignificant.

Using the individual-state sales change, we continue to find a positive and significant coefficient on the New Year dummy, with its magnitude implying an increase in sales of approximately \$120,000 per half a week within each state during early January. In untabulated tests, we split the sale changes scaled by the size of the jackpot between January and other months and find that the adjusted sales change is always greater in January than in other months across all years in our sample. In short, the lottery sales data provide support for a stronger gambling mentality at the turn of the New Year. Since regular casino gamblers and typical lottery players possess similar demographic profiles to investors that prefer lottery-type stocks (Kumar, 2009), we expect that the mentality of casino gamblers reflects that of lottery-type asset players in the financial markets.

2.4 HYPOTHESES

Based on the existing literature reviewed above and our direct evidence from gaming data, we expect securities with lottery features will outperform in January as a result of investor excess demand. We lay out the testable hypotheses below.

In the US option markets, we hypothesize that investors demand OTM more than ATM call options in January to a greater extent than during other months of the year. This is due to the lottery feature associated with OTM calls. We assess the price impact through measuring the implied volatility spread between OTM and ATM calls. We assess the volume impact through the trading volume spread between the two. In addition, we expect a stronger shift toward lottery-type bets among small investors (customers) than large institutions (firms), as evidenced by the purchase of more calls relative to puts in January than in other months.

Hypothesis 1. *In the option markets, (a) The implied volatility spread and trading volume spread between OTM and ATM call options should be greater in January than in other months. (b) Individual investors, particularly small ones, should have higher call/put open buy ratios than institutions in January than in other months.*

In the US stock market, we examine the cross-sectional stock return patterns across lottery features between January and other months.

Hypothesis 2. *Stocks with strong lottery features (low price, high idiosyncratic volatility, and high idiosyncratic skewness) should outperform those with weak lottery features in January but not necessarily in other months.*

We further examine whether retail trades in equity markets exhibit seasonal preferences toward lottery-type stocks.

Hypothesis 3. *In the US stock markets, (a) Retail investors should be more bullish toward lottery-type stocks in (early) January than in other times of the year. (b) The outperformance of lottery-type stocks should be stronger for those with more bullish retail sentiment or more actively traded by retail investors. (c) The outperformance of lottery-type stocks should be stronger in trading days surrounding New Year's Day than other trading days later in January.*

Hypothesis 3a links the seasonality in retail sentiment to that of the lottery-type stocks. As a result, in Hypothesis 3b, we predict that stocks more prone to the swing of retail sentiment should experience greater price appreciation in January. Since such gambling sentiment is likely to be stronger surrounding New Year's Day, we further predict in Hypothesis 3c that the price run-up of lottery stocks should concentrate around New Year's Day.

In addition, we test the gambling preference–based hypothesis against traditional hypotheses involving individual tax-loss selling, institutional window-dressing, or institutional risk shifting. These traditional hypotheses predict that the January effect of lottery-type stocks should occur solely for past losers or for stocks bought by institutions. In contrast, the gambling preference–based hypothesis predicts that it should occur regardless of past returns and institutional trading. We also study the return performance of lottery stocks purchased by institutions versus individual investors. We expect that the purchase of lottery stocks by individuals is mostly driven by a gambling preference as opposed to a stock selection skill. Thus, lottery-type stocks selected by individuals should have lower expected returns than those chosen by institutions.

In the China stock markets, we focus on the comparison between Chinese New Year's months and January trading days that precede the Chinese New Year's Day. We examine the market as a whole and the relative performance between lottery-type Chinese stocks and their counterparts.

Hypothesis 4. *In the China stock markets, (a) The market as a whole should outperform in the Chinese New Year's Month, but not necessarily in January, relative other months. (b) Chinese stocks with strong lottery features should outperform those with weak lottery features in the Chinese New Year's Month but not necessarily in January.*

3. Data and Measures of Lottery Features

3.1 DATA

Our US stock sample includes all common stocks (share codes 10 and 11) listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) from July 1963 through December 2007. Stock returns and other trading data are from the Center for Research in Securities Prices (CRSP). Accounting information for the calculation of book equity is from COMPUSTAT. Implied volatility from call options and option volume are from OptionMetrics from 1996 to 2006. Aggregate weekly call and put volume for equity-based options across different categories of investors (customer, firm, and market maker) are from the Option Clearing Corporation (OCC), available from 2000 to 2008. The retail and institutional orders are inferred from the tick-by-tick transaction data compiled by ISSM from 1983 to 1992 and TAQ data from 1993 to 2000 for NYSE and AMEX common stocks.[17] Chinese stock market data are from DataStream over the period 1993–2006. The Fama and French factor returns are from Kenneth French's Web site. Institutional ownership is from the Thomas Financial 13F file from 1980 to 2007.

3.2 US OPTION MONEYNESS AND OPEN BUY VOLUME

OptionMetrics gives daily information about implied volatility and trading volume for all tradable options. We define OTM (ATM) calls as those with the ratio of the strike price to the stock price greater than 1.05 (between 0.975 and 1.025). Implied volatility is measured on the last trading day of each month for options that expire in the following month. We compute the average implied volatility separately for all OTM and ATM calls for each stock. Then, we define the monthly implied volatility spread (OTM − ATM) as the mean difference between OTM and ATM implied

[17] Following Barber, Odean, and Zhu (2009), we end in 2000 when decimalization is widely introduced, making it difficult to distinguish institutional from retail trades based on the size of the order. We only have access to NASDAQ data from 1993 to 2000. Our results are robust to including NASDAQ data for this period.

volatilities across all stocks in a given month. For a firm's options to be included, they must have positive trading volume for at least one ATM and one OTM call option for contracts expiring the following month.

For each call option, monthly volume is the sum of daily volume. Aggregate monthly volume for OTM (ATM) calls is the total monthly trading volume across all OTM (ATM) calls in a given month. Since average option volume has an upward time trend, we separately compute for OTM and ATM calls the percentage change of the current month option volume from its past 12-month moving average and label it as adjusted option volume. Then, we calculate the monthly adjusted volume spread (OTM − ATM) as the average difference between the adjusted option volume of OTM and ATM calls.[18]

Weekly OCC volume reports provide the aggregate number of buy and sell contracts for all equity-based options across all exchanges, strike prices, and maturities for each account type. Moreover, these reports separate open call and put buys and sells by firm, customer, and market maker. We focus on the first two account types, where a "firm" account is established by a clearing member for its purposes, and a "customer" account is established by a clearing member on behalf of its customers. The clearing members tend to be large institutions, while customers include retail investors and nonclearing member institutions (such as small hedge funds).

Based on the size of each transaction, customers are further classified into three groups: 1–10 contracts per transaction, 11–49, and above 50. Following Battalio, Hatch, and Jennings (2004), we use the group with 1–10 contracts per transaction, which we refer to as small customers, as a proxy for retail orders. Based on whether the Friday of a week is in January, we classify the weeks as January and non-January. We are interested in whether more open calls, relative to open puts, are bought in January weeks for all types of investors, and whether this open buy call/put ratio is highest for customers, particularly for small customers versus firms.

3.3 US STOCK LOTTERY FEATURE MEASURES

Following Kumar (2009) and Boyer, Mitton, and Vorkink (2010), we use three measures of stock lottery features. The first is stock price (PRC), the closing price at the end of the immediately preceding month. Like lotteries, stocks with low prices attract gamblers. The second is idiosyncratic volatility (IVOL). High idiosyncratic volatility inflates the perception of the chance to realize high returns, thus

[18] A concern is that for a given month, the OTM options have a different strike to stock ratio than other months, resulting in comparing the implied volatility or volume at different points on the volatility skew. However, we find that the average OTM strike/spot ratio for the sample is 1.17, with no significant differences across months. The lowest is 1.16 (August) and the highest is 1.19 (April). For ATM strike/spot ratios, the average is 0.999.

attracting stock market gamblers. Following Ang *et al.* (2006), IVOL is the standard deviation of at least seventeen daily residual returns within the preceding month, where residual returns are estimated from the Fama and French (1993) three-factor model. Our main findings for January hold for a number of alternative measures of idiosyncratic volatility, including that estimated from either the past 3-month daily or the 12-month weekly returns, inferred from option prices and estimated from an Exponential Generalized AutoRegressive Conditional Heteroskedasticity model. Section B of the online Supplementary Appendix presents the measures and results.

The third lottery feature is expected idiosyncratic skewness. Boyer, Mitton, and Vorkink (2010) show that future skewness is poorly predicted from past skewness; thus, it is important to estimate expected, not realized, idiosyncratic skewness. We employ the approach of Boyer et al. that uses a set of firm characteristics to forecast future idiosyncratic skewness through monthly cross-sectional regressions.[19]

Specifically, for each individual stock, we first calculate realized idiosyncratic skewness (ISKEW) based on at least twenty-six of fifty-two weekly residual returns over a rolling 12-month window. The residual returns are estimated from regressions of weekly stock returns on weekly market returns and squared weekly market returns (Harvey and Siddique, 2000; Kumar, 2009).[20] Then, we run month-by-month cross-sectional firm-level regressions. The dependent variable is the future 12-month realized ISKEW measured from month t. The independent variables include the past 12-month ISKEW value and a set of other firm characteristics suggested by Boyer, Mitton, and Vorkink (2010). To obtain the expected idiosyncratic skewness (EISKEW), we apply the regression coefficients estimated in month $t-12$ on independent variables observed in month $t-1$. For example, in December 1997, we use the coefficients estimated in December 1996 and the skewness predictors measured as of December 1997 to compute EISKEW. Then, we use EISKEW to forecast stock returns in January 1998. In other words, EISKEW is computed with prior information and used to forecast future returns. Our main results are insensitive to alternative specifications to forecast EISKEW. Section C of the online Supplementary Appendix presents robustness checks based on five alternative specifications.

[19] In unreported tests, we use realized idiosyncratic skewness, defined following Kumar (2009), to assess robustness of our results. For both equal- and value-weighted portfolio returns, we still find that highly skewed stocks significantly outperform their counterparts in January. We do not observe significant underperformance of highly skewed stocks in non-January months, consistent with the argument of Boyer, Mitton, and Vorkink (2010) that past skewness is a poor predictor of future skewness and thus unlikely to capture the negative premium of high skewness stocks.

[20] Our results hold if the residual returns are estimated from the Fama-French three-factor model, and if the rolling window is three or six months long. We focus on the 12-month measures because our hypothesis implies that investors buy lottery-type stocks in January for a chance of a “home run” in the year.

Although our main findings hold for each of the lottery features, we also construct a lottery-feature index (LOTT) that incorporates PRC, IVOL, and EISKEW. To construct LOTT, we independently rank all stocks into twenty groups for each lottery measure. Then, we assign a ranking to each of the three sets of portfolios with the highest ranking, 19, assigned to the lowest PRC, highest IVOL, and highest EISKEW portfolios and the lowest ranking, 0, assigned to the other three extreme portfolios. The portfolio rankings are assigned to each stock included in the portfolio. For each stock, we define the average of the three rankings as the lottery feature index LOTT.[21]

3.4 RETAIL AND INSTITUTIONAL ORDERS

To infer retail and institutional orders from tick-by-tick transactions, we follow the procedure by Barber, Odean, and Zhu (2009). First, we identify trades as buyer or seller initiated, following Lee and Ready (1991), based on the tick and the trade rules. Since we use NYSE/AMEX stocks, we exclude the opening trades because they represent a call auction. Then, for trades executed at the midpoint of the bid and ask price, we apply the tick rule before the quote rule, and for others, the quote rule before the tick rule. Under the quote rule, trades are buyer initiated if the trade price is above the midpoint of the most recent big-ask quote and seller initiated if the trade price is below the midpoint. Under the tick rule, trades are buyer initiated if the trade price is above the last executed trade price and seller initiated if the trade price is below it.

Again following Barber, Odean, and Zhu (2009), trades at or below \$5,000 in 1991 are classified as retail trades and those above \$50,000 in 1991 are classified as institutional trades. In other years, we adjust the cutoffs according to the consumer price index. We identify retail and institutional buy and sell orders on a daily basis. We compute the retail trading proportion measure (RTP), following Brandt *et al.* (2010), as the total dollar amount of retail trades over the total dollar amount of all trades in a given month. Following Barber, Odean, and Zhu (2009), we compute the retail order imbalance (RIMB), defined as the dollar amount of buyer- minus seller-initiated trades over the total dollar amount of buyer- and seller-initiated trades for a given day or month. This measure captures the relative bullish sentiment of retail investors about a given stock. To compare it with the institutional sentiment, we also compute the institutional order imbalance (IIMB), which is defined similarly to RIMB but over institutional orders.

[21] To maximize the number of observations of lottery-type stocks, we only require one of the three rankings to be available to compute LOTT. However, our results hold if we require two or three rankings to compute LOTT.

3.5 CHINA STOCK LOTTERY FEATURE MEASURES

For the China stock sample, we measure lottery features with logarithmic price (LOGPRC) and IVOL. We forgo expected idiosyncratic skewness because many firm characteristic variables are unavailable. Stock price, in local currency, is measured at the end of the preceding month. IVOL is the standard deviation of residual returns from a regression of at least fifteen daily returns in the preceding month on contemporaneous daily returns on the Shanghai and the Shenzhen (two major exchanges in China) index returns. Sometimes in January and March, the number of trading days for the overall market is less than 15. This occurs when the Chinese New Year's Day is in early January or late February. If this happens, IVOL of month $t-2$ forecasts returns for month t. Stock returns are measured in local currency.

Since the Chinese New Year's Day can occur anywhere in January or February, we define January, the Chinese New Year's Month, and March as including January trading days prior to the Chinese New Year's Day, the 22-trading-day period following the Chinese New Year's Day, and the subsequent trading days through the end of March, respectively.[22] We cumulate returns over each of the redefined months and normalize them to a 22-day return. We also compute monthly returns from April to December. Thus, all returns are expressed on a monthly basis. The equal-weighted monthly return across all available stocks is a proxy for the equal-weighted market portfolio.

3.6 SUMMARY STATISTICS

The US option sample is described in Panel A of Table II. The average annualized implied volatility of OTM calls is 55.32% higher than that for ATM calls of 42.33%. Implied volatility for OTM calls greater than that for ATM calls is consistent with the smile behavior for individual equity options versus index options (Bakshi and Kapadia, 2003). The ratio of call to put buys is larger for individual customer open buy transactions than firm open buy transactions and especially so for the smaller customer open buy transactions. The volume and call/put ratios are generally consistent with Pan and Poteshman (2006).

Panel B of Table II reports the summary statistics for the lottery feature measures. The average stock price is roughly \$10 ($\approx e^{2.27}$), the average daily idiosyncratic volatility is 2.96% or 47% on an annualized basis, idiosyncratic weekly skewness

[22] Since the Chinese New Year's Day tends to occur somewhere from mid-January to mid-February, our definition classifies the trading days in February preceding the Chinese New Year's Day as belonging to no group. However, our results are similar if we classify those trading days as if they occur in March, which causes the trading days defined in March to be skipped by the New Year's Month in some years.

Table II. Summary statistics for US option, US stock, and China stock samples

The first two rows of Panel A are summary statistics for the US option sample from 1996 to 2006. The average annualized implied volatility (ImpVol), monthly volume (Volume), and adjusted option volume (Adj. volume) across all firm-months are reported. OTM (ATM) calls have the ratio of the strike price to the stock price greater than 1.05 (between 0.975 and 1.025). Only options expiring in the following month and with nonzero trading volume are included. For a given stock, at the end of each month, we compute the implied volatility as the average annualized implied volatility (in %) of all OTM (ATM) calls. For a given stock, the monthly volume of OTM (ATM) calls is calculated as the average monthly trading volume across all OTM (ATM) calls in a given month. The adjusted volume is the percentage change of the current month volume from its past 12-month moving average to account for the upward time trend in option volume. The last five rows of Panel A show equity-based option volume data from OCC reports, including all strikes and maturities from 2000 to 2008. The average weekly option volume (Volume), call volume (Call), put volume (Put), and C/P ratios (C/P) across all account types are reported. For a given stock, weekly volume refers to the average weekly option volume, call volume refers to the average weekly volume on open buy call options, and put volume refers to the average weekly volume on open buy put options, all in thousands of contracts. The C/P ratio is the ratio of open buy call option volume divided by open buy put option volume. Open buys are new purchases of call options and include no buying to cover previous short call positions. ‘‘Firm’’ refers to an account established by a clearing member for the purposes of the clearing member. ‘‘Customer’’ refers to an account established by a clearing member on behalf of its customers. Based on the size of each transaction, customers are classified into three groups: 1–10 contracts per transaction, 11–49, and 50 and above. Panel B reports summary statistics of the US stock sample over the period July 1963 to December 2007. The lottery feature measures include stock price (PRC) (or logarithmic stock price, LOGPRC), idiosyncratic volatility (IVOL), idiosyncratic skewness (ISKEW), expected idiosyncratic skewness (EISKEW), and the lottery feature index (LOTT). PRC is the closing stock price measured at the end of the preceding month. IVOL is the standard deviation of at least seventeen daily residual returns from the Fama–French three-factor model within the preceding month. ISKEW is the skewness of at least 26 weekly residual returns over the past 12 months, where the residual returns are obtained from the model with the market and squared market factors. EISKEW is the expected 12-month ahead weekly idiosyncratic skewness based on monthly forecasts. LOTT is the average ranking of PRC, IVOL, and EISKEW. For each feature, the ranking is based on 20 portfolios sorted by the feature. These portfolios are indexed from 0 to 19, with 19 indicating the lowest PRC, highest IVOL, or highest EISKEW group. Panel C reports the summary statistics of the China stock sample from 1994 to 2006 for the monthly return (RET), LOGPRC, and IVOL. RET refers to percentage monthly returns, with months redefined to account for the Chinese New Year. LOGPRC is in local currency measured at the end of the preceding month. IVOL is the standard deviation of at least 15 daily residual returns in the preceding month from the regression of daily returns on both the Shanghai and Shenzhen A share indices.

Panel A: US option sample				
OptionMetrics	Average no. of firms	ImpVol	Volume	Adj. volume
ATM calls	1,963	42.33	1,852	0.119
OTM calls	1,963	55.32	775	0.077

(*Continued*)

Table II. (Continued)

OCC reports	Volume	Call	Put	C/P
Firm	3,684	1,885	1,799	1.30
Customer	11,542	6,617	4,925	1.64
1–10	2,149	1,404	745	2.23
11–49	2,320	1,448	873	2.05
50 and above	7,073	3,765	3,307	1.36

Panel B: US stock sample

	No. of firm-month observations	Mean	Median	Standard deviation
LOGPRC	2,545,129	2.27	2.50	1.25
IVOL	2,513,228	2.96	2.23	2.75
ISKEW	2,471,403	0.62	0.52	1.02
EISKEW	1,868,909	0.60	0.58	0.38
LOTT	2,558,395	9.66	10.00	4.78

Panel C: China stock sample

	No. of firm-month observations	Mean	Median	Standard deviation
RET	72,376	0.21	−1.06	14.70
LOGPRC	72,376	2.01	2.03	0.54
IVOL	72,376	1.61	1.42	0.87

is 0.62, and expected weekly idiosyncratic skewness is 0.60. Panel C describes the China stock sample. The average monthly stock return is 0.21% with a negative median return (−1.06%) and a large standard deviation (14.70%). This is because there are predominately more firms in the latter period of the sample during which China stock markets have poor performance. In our subsequent tests, we use cross-sectional demeaned stock returns as dependent variables. Thus, our results are not driven by certain unique periods in the sample.

Panel A of Table II presents characteristics of quintiles of stocks sorted on LOTT. The average log price is 3.57 (≈\$35 per share) for the lowest LOTT quintile of stocks and 1.06 (≈\$2.88 per share) for the highest quintile. Although there are roughly equal numbers of firms (about 1000) in each quintile, their market capitalizations proportional to the overall market are quite different. The lowest LOTT quintile accounts for more than 80% of the total market cap, while the highest LOTT quintile accounts for merely 1%. Our subsequent tests show that the relation between lottery features and January returns is monotonic. Thus, even after we exclude the highest LOTT quintile of stocks, there still exists significant return differentials across LOTT groups. More importantly, we find that the RTP is highest for the highest LOTT quintile, for which 26% trades are retail, as opposed to only 1% for the lowest LOTT quintile. This evidence suggests that retail investor

Table III. Summary statistics of lottery feature portfolios, correlation matrix, and regressions to forecast idiosyncratic skewness

Panel A reports the characteristics of quintiles sorted on the lottery feature index (LOTT) over the period July 1963 to December 2007. LOGPRC is the logarithmic stock price, IVOL is the idiosyncratic volatility, EISKEW is the expected idiosyncratic skewness, and ISKEW is the idiosyncratic skewness. All are defined in Table II. ME (%) refers to the total market capitalization within the quintile as a percentage of the total market cap across all firms. RTP is the RTP, defined as the total dollar amount of retail trades over the total dollar amount of all trades in a given month. Panel B reports the correlations between lottery feature measures and other firm characteristics. RET(−1) and RET(−12, −2) refer to the return in the prior month and the prior 2 through 12 months, respectively. LOGME is logarithmic firm size, defined as the market equity measured at the end of the preceding month. LOGBM is logarithmic book-to-market equity. Book-to-market equity, from June of year s through May of year $s + 1$, is book equity measured at the fiscal year-end through December of year $s - 1$, over market equity measured at the end of December of year $s - 1$. Share turnover (TURN) is the total monthly trading volume over shares outstanding in the preceding month multiplied by 100. Trading volume of NASDAQ firms is divided by two. Panel C reports the average coefficients of the monthly cross-sectional regressions of future ISKEW measured over months t through $t + 11$ on past ISKEW measured over months $t - 12$ through $t - 1$ and a set of firm characteristics. EISKEW in month t is calculated based on ISKEW and other firm characteristics observed at the end of month $t - 1$ and the regression coefficients estimated in month $t - 12$. The time series means of the coefficients from monthly cross-sectional regressions are reported with the Newey-West (1987) t statistics.

Panel A: summary statistics of portfolios sorted based on LOTT

LOTT rank	Average no. of firms	LOTT	LOGPRC	IVOL	EISKEW	ISKEW	ME (%)	RTP
L	958	2.85	3.57	1.26	0.25	0.28	81.86	0.01
2	964	6.70	2.87	1.79	0.47	0.50	10.75	0.02
3	959	9.80	2.46	2.22	0.63	0.63	4.25	0.04
4	954	12.83	1.83	3.11	0.80	0.77	2.01	0.09
H	949	16.13	1.07	5.58	1.05	1.00	1.14	0.26

Panel B: correlation matrix

	IVOL	ISKEW	EISKEW	LOTT	LOGME	LOGBM	RET (−1)	RET (−12, −2)	TURN
LOGPRC	−0.54	−0.19	−0.80	−0.75	0.73	−0.09	0.09	0.24	0.08
IVOL		0.15	0.51	0.58	−0.34	−0.06	0.15	−0.12	0.12
ISKEW			0.37	0.25	−0.24	0.09	0.11	0.18	−0.01
EISKEW				0.82	−0.83	0.18	−0.07	−0.27	−0.19
LOTT					−0.73	0.05	−0.01	−0.16	−0.04

Panel C: firm-level monthly cross-sectional regression to forecast EISKEW

	Intercept	LOGPRC	IVOL ×100	ISKEW ×100	LOGME ×100	LOGBM ×100	RET (−1) ×100	RET (−12, −2) ×100	TURN ×100
Coefficient	1.22	−0.11	1.03	0.05	−0.08	1.60	−0.10	−0.09	−0.30
t Statistic	61.42	−16.48	6.31	26.97	−26.98	5.94	−7.90	−14.09	−6.25

trading behavior and sentiment should have larger impacts on stocks with the most salient lottery features than those with the least.

The correlation matrix between lottery feature variables and firm characteristics are reported in Panel B. The correlation matrix in Panel B shows that LOGPRC, IVOL, and EISKEW are highly correlated, but the correlations between past skewness (ISKEW) and other lottery features are much weaker. This is consistent with prior literature showing that expected skewness more accurately represents stock lottery features than past skewness.

Panel C reports the Fama and MacBeth (1973) regression coefficients that forecast EISKEW. The independent variables include LOGPRC, IVOL, ISKEW, logarithmic size (LOGME), logarithmic book-to-market equity (LOGBM), past 1-month returns (RET(-1)), past returns from month $t-12$ through $t-2$ (RET($-12, -2$)), and share turnover in the previous month (TURN), all defined in Table III. The forecast regression requires that all predictors are available. Thus, the total number of observations for which we have EISKEW available are reduced by about 25% from those with ISKEW.[23] All coefficients are statistically significant and their signs are consistent with those in Boyer, Mitton, and Vorkink (2010).

4. Empirical Results

4.1 US OPTION MARKETS

4.1.a. Implied volatility and volume spreads

For our option sample, we examine option pricing and trading. We test Hypothesis 1a that the implied volatility spread and trading volume spread between OTM and ATM calls should be greater in January than other months.

Panel A of Table IV reports the average monthly implied volatility spread of OTM and ATM calls between January and non-January months from 1996 to 2006. We also report the maximum and minimum values between February and December. Consistent with our conjecture, the implied volatility spread is higher in January than all other months, suggesting that OTM calls are relatively more expensive in the New Year. The average difference of the annualized implied volatility spread between January and other months is 4% and statistically significant ($t = 5.55$).

Panel B of Table IV reports the average adjusted option volume spread between January and non-January months, including the maximum and minimum average

[23] Depending on the specification, the reduction of the number of available firms can be less. Section C of the online Supplementary Appendix shows alternative specifications that require fewer predictors and, hence, have more firm-month observations. Our main results hold for these samples too.

Table IV. Implied volatility and adjusted option volume of call options

Panel A gives the mean monthly implied volatility spreads, first averaged across all firms and then averaged across years for a given month, separately reported for January and non-January months. It shows that the implied volatility spread between OTM and ATM calls is higher in January than in other months. That is, relative to ATM calls, OTM calls are the most expensive in January. Panel B gives the average monthly adjusted volume spread, first averaged across all firms and then averaged across years for a given month, separately reported for January and non-January months. It shows that the adjusted option trading volume spread between OTM and ATM calls is higher in January than in other months. That is, relative to ATM calls, OTM calls are most heavily traded in January. OTM and ATM calls are defined in Table II. The maximum and minimum values in both Panels A and B are based on the average values for each month from February through December. Panel C reports the mean open buy call/put ratios by customer type and firm. The difference between January and non-January is reported under Jan − (NonJan). The difference between customer type and firm of open buy call/put ratios in January is reported for each customer type under $(\text{Customer} - \text{firm})_{\text{Jan}}$. In Panel C, January weeks are defined as all weeks with a Friday ending in January. Non-January weeks include all other weeks. The difference between customer type and firm of open buy premium differentials, between January and other months, is reported for each customer type under $(\text{Customer} - \text{firm})_{\text{Jan} - \text{NonJan}}$.

Panel A: implied volatility

	Mean		Minimum	Maximum
	January	Non-January	Non-January	
OTM	57.36	54.40	51.92	57.00
ATM	43.88	45.01	42.80	48.09
OTM − ATM	0.13	0.09	0.07	0.11
t(OTM − ATM)	8.00	13.61		
Jan − (NonJan)	0.04			
t[Jan − (NonJan)]	5.55			

Panel B: option volume

	Mean		Minimum	Maximum
	January	Non-January	Non-January	
OTM	0.49	0.04	−0.27	0.34
ATM	0.27	0.10	−0.21	0.40
OTM − ATM	0.22	−0.06	−0.37	0.13
t(OTM − ATM)	1.79	−1.49		
Jan − (NonJan)	0.29			
t[Jan − (NonJan)]	2.09			

Panel C: open buy call/put ratio

	Customer				Firm
	All	1–10	11–49	50 and above	All
Jan	1.613	2.429	2.060	1.342	1.293
NonJan	1.451	2.122	1.838	1.226	1.165
Jan − (NonJan)	0.162	0.307	0.222	0.116	0.128
t[Jan − (NonJan)]	1.89	2.04	2.19	1.90	2.39
$(\text{Customer} - \text{firm})_{\text{Jan-(NonJan)}}$	0.320	1.140	0.770	0.050	
$t[(\text{Customer} - \text{firm})_{\text{Jan-NonJan}}]$	2.21	5.50	4.62	0.49	

values from February to December. As expected, the adjusted volume spread is higher in January than other months. The average difference in the adjusted volume spread is 0.29, which is statistically significant ($t = 2.09$). Relative to ATM calls, OTM calls are traded most in January. Thus, our evidence supports Hypothesis 1a that there is excess demand for OTM calls at the turn of the year.

4.1.b. OCC open buy call/put ratios and call and put premiums

To test Hypothesis 1b, we examine the seasonality in customers' call/put open buy ratios versus those for firms. Panel C of Table IV shows the average call/put open buy ratios for January and all other months. The results show two findings consistent with our hypothesis. First, for both customers and firms, the open buys of calls, relative to puts, is higher in January than other months. Second, the average customer open buy call/put ratio is significantly higher than that of the firm in January. This is especially true for small customer purchases of between one and ten contracts. The difference in the ratios is 1.14 with a t statistic of 5.50. While both firms and customers purchase more calls in January, smaller customers are much more inclined to do so. Overall, the results provide direct support for our hypothesis that investors demand lottery-type options and impact option prices and volume.

4.2 US STOCK SAMPLE

4.2.a. Sorts

We now test Hypothesis 2 about whether lottery-type stocks outperform those with opposite characteristics mainly in January.

Each month we independently sort stocks into quintiles based on the three lottery feature variables PRC, IVOL, and EISKEW and the lottery feature index LOTT. We then compute the average value- and equal-weighted returns of each quintile for the subsequent month. We form hedge portfolios ($S - W$) that are long the strongest lottery feature (low PRC, high IVOL, EISKEW, and LOTT) and short the weakest lottery feature quintiles and compute mean returns and alphas from the Fama and French (1993) three-factor model for these portfolios. Our results are similar if we add the momentum factor to the three-factor model. We start by examining the relation between lottery features and stock returns across all months. The purpose is to replicate prior results using the more recent sample and set the benchmark for separating January analyses from other months.

As shown in Table V, the average relation between lottery features and stock returns across all months is mixed. With value-weighted portfolio returns, we find

Table V. Monthly returns of portfolios sorted on lottery features

This table shows that stocks with strong lottery features outperform those with weak lottery features in January but tend to underperform in other months for both value-weighted (Panel A) and equal-weighted (Panel B) returns. It reports the average monthly percentage returns of quintiles of stocks sorted on stock price (PRC), idiosyncratic volatility (IVOL), expected idiosyncratic skewness (EISKEW), and the lottery feature index (LOTT). The variables PRC, IVOL, EISKEW, and LOTT are defined in Table II. Average percentage returns of these quintiles, from July 1963 to December 2007, are shown across all months, in January and in non-January months. W (weak) refers to the lowest LOTT, IVOL, or EISKEW and the highest PRC quintile. S (strong) refers to the highest LOTT, IVOL, or EISKEW quintile and the lowest PRC quintile. $S-W$ refers to the long minus short portfolio that is long the strongest lottery feature quintile and short the weakest quintile. $\alpha(S-W)$ for all months, January, and non-January months refers to the monthly average abnormal return across those periods, where the abnormal return is defined relative to the three-factor model with factor loadings estimated using all months over the full sample. The Newey-West (1987) t statistics are reported in parentheses.

	All months				January				Non-January			
Rank	LOTT	PRC	IVOL	EISKEW	LOTT	PRC	IVOL	EISKEW	LOTT	PRC	IVOL	EISKEW
					Panel A: value weighted							
W (weak)	0.96	0.94	0.96	0.89	1.61	1.51	1.53	1.46	0.90	0.89	0.91	0.83
2	1.03	1.03	1.03	1.07	3.52	2.97	2.19	2.91	0.80	0.86	0.92	0.90
3	0.82	1.00	1.07	1.05	3.67	4.54	3.08	4.62	0.56	0.69	0.89	0.73
4	0.59	0.83	0.73	0.92	5.88	7.55	3.65	7.63	0.12	0.22	0.47	0.31
S (strong)	0.38	0.90	−0.01	1.04	8.82	12.56	5.43	12.36	−0.38	−0.15	−0.50	0.01
$S-W$	−0.58	−0.04	−0.97	0.15	7.21	11.05	3.91	10.90	−1.28	−1.04	−1.41	−0.83
	(−1.74)	(−0.11)	(−3.02)	(−0.44)	(4.04)	(5.47)	(3.41)	(5.44)	(−3.80)	(−3.07)	(−4.19)	(−2.45)
$\alpha(S-W)$	−0.86	−0.63	−1.26	−0.46	4.13	6.53	0.92	5.94	−1.31	−1.28	−1.46	−1.05
	(−4.09)	(−2.69)	(−6.38)	(−2.32)	(2.60)	(4.02)	(1.59)	(3.92)	(−5.87)	(−5.27)	(−6.99)	(−4.94)
					Panel B: equal weighted							
W (weak)	1.20	1.14	1.17	1.04	2.24	1.90	2.83	1.72	1.11	1.08	1.02	0.98
2	1.33	1.20	1.39	1.17	3.99	3.36	4.15	3.04	1.09	1.01	1.14	1.00
3	1.24	1.16	1.40	1.20	5.16	5.04	5.60	4.96	0.89	0.81	1.02	0.86
4	1.19	0.94	1.25	1.07	8.03	7.75	7.71	7.83	0.58	0.32	0.67	0.45
S (strong)	1.42	1.94	1.10	1.92	13.72	15.05	12.85	14.80	0.32	0.76	0.04	0.75
$S-W$	0.22	0.79	−0.07	0.88	11.48	13.15	10.01	13.08	−0.80	−0.32	−0.98	−0.23
	(−0.69)	(−2.50)	(−0.23)	(−2.81)	(5.51)	(6.40)	(5.63)	(6.10)	(−2.64)	(−1.12)	(−3.03)	(−0.82)
$\alpha(S-W)$	−0.08	0.36	−0.39	0.37	8.51	9.76	7.05	8.98	−0.86	−0.49	−1.07	−0.43
	(−0.37)	(−1.53)	(−1.87)	(−1.72)	(4.35)	(5.05)	(4.60)	(4.70)	(−4.04)	(−2.13)	(−5.05)	(−1.99)

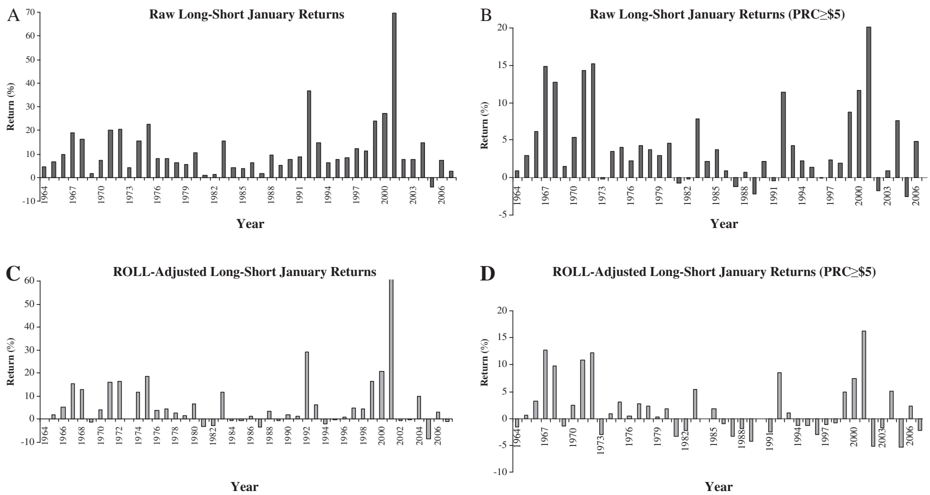

Figure 2. January returns on the hedge portfolios based on the lottery feature index. Panels A and B depict the year-by-year January raw returns on the hedge portfolio that is long the highest lottery index (LOTT) quintile and short the lowest quintile. In Panel B, stocks with an end-of-December price below \$5 are excluded. Panels C and D depict the ROLL-adjusted January returns for the two portfolios. The returns are equal-weighted. For each firm we compute the bid–ask spread–adjusted returns as the raw percentage returns minus the percentage effective bid–ask spread (ROLL) proposed by Roll (1984). These figures show that the trading strategy based on lottery features involves limited downside risk over the sample period.

that low PRC, high IVOL, and high LOTT stocks tend to underperform their counterparts, controlling for the Fama–French three factors. The highest IVOL quintile underperforms the lowest by 0.97% per month ($t = -3.02$) similar in magnitude to the 1.06% found by Ang *et al.* (2006) from 1963 to 2000. Using equal-weighted returns, however, this relative underperformance diminishes, consistent with Bali and Cakici (2008). The returns on the hedge portfolios are mixed in sign, and none of the alphas are statistically significant at the 5% level. Across all months, whether lottery-type stocks under- or overperform is sensitive to the portfolio weighting scheme.

We next separate January returns from those of other months. Hypothesis 2 predicts that there should be a positive relation between lottery features and stock returns in January. The monthly separation shows that January returns across lottery features differ sharply from the non-January return patterns. For all four lottery features, and for both equal- and value-weighted returns, there is a strong positive relation between lottery features and stock returns. The hedge portfolios based on PRC, IVOL, ESIKEW, and LOTT produce average January returns of 3.91–13.15%. All January alphas but one are statistically significant at the 1% level. In contrast, in non-January months, there is a consistent negative relation between lottery features and portfolio returns across all lottery features. All alphas are negative and statistically significant at the 5% level or better.

In other words, there are opposite relations between lottery features and portfolio returns across January and non-January months. The positive relation in January is consistent with our hypothesis that investor preference for lottery-type stocks impacts returns at the turn of the year. The negative relation in non-January months, consistent with behavioral models (Mitton and Vorkink, 2007; Barberis and Huang, 2008; Barberis and Xiong, 2008), suggests a correction of overpricing of lottery-type stocks caused by investor preference for speculative features. Our findings also show that the January effect of lottery-type stocks is responsible for the difference in return patterns between equal- and value-weighted portfolios. If Januarys are excluded, there is a consistent negative relation between IVOL/EISKEW and stock returns regardless of the weighting scheme.

In unreported results, we also study the month-by-month long-short portfolio returns from January through December to check possible quarter-end effects. We find that lottery-type stocks significantly outperform in January, slightly outperform in February, and tend to underperform in the remaining months. These results suggest that the January effect of lottery stocks is unlikely to be caused by institutional risk-shifting behavior at the end of the quarter.

4.2.b. The January trading strategy

Is the lottery-stock-in-January effect exploitable? We examine the implication of our results for trading strategies for two reasons. First, it is important to understand whether our results have value for practitioners. Second, it is useful to see whether any trading profits survive microstructure considerations, such as the bid–ask spread.

We consider the equal-weighted hedge portfolios based on the LOTT quintiles, formed at the end of each December and liquidated at the end of the following January, since we have shown that the equal-weighted hedge portfolios deliver higher expected returns than value-weighted ones. Thus, it should be a more attractive investment strategy.

In Panel A of Figure 2, we present an annual display of the January returns over the 44 years from 1964 through 2007. The mean raw return of 10.40% per month is earned with only one yearly loss in 2005 (−6.04%). The highest profit year is 2001 with a remarkable 72% January return. Overall, high returns seem to come with limited downside risk.[24] An investor who starts in 1965 with $1 in the hedge portfolio and invests from the end-of-December to the end-of-January would end up

[24] The impact of short-sale constraints is relatively minor here because the short position contains high price (mostly large firms) and low volatility stocks, which are generally easy and cheap to sell short.

with over \$92.84 by the end of January 2007. By contrast, those who trade in the opposite direction would virtually lose all their investment.

To see if profits are robust to the adjustment of trading costs, we consider three factors. First, we consider the bid–ask spread in trading the long and short portfolios. Keim (1989) suggests that there is a systematic shift of the closing price from the bid in December to the ask in January, artificially inflating returns, particularly on low-price stocks. Thus, it is important to assess whether the outperformance of lottery-type stocks persists after accounting for the spread. We use the Roll (1984) method to compute the effective percentage bid–ask spread (ROLL), defined as the square root of the negative autocovariance of weekly returns from February to November of the prior year, multiplied by 200. Our estimate of the mean ROLL is 2.79% for the lowest LOTT quintile and 8.02% for the highest LOTT quintile.[25] Panel C displays the equal-weighted ROLL-adjusted January returns for the long-short LOTT portfolio. The spread-adjusted returns are smaller than the raw returns, but the superior performance of high LOTT stocks remains, with relatively limited risk. Few years have negative returns and when they occur, the losses tend to be small. The average net return is 6.42% ($t = 3.69$), which remains economically significant.

Second, we eliminate stocks below \$5 since institutions usually find it unattractive to trade these stocks due to large indirect costs (Keim and Madhavan, 1998). Panels B and D plot the annual January returns for the long-short LOTT portfolio that is implemented among all stocks and those trading for \$5 or more. Compared to before, there are a few more years with negative returns, but they tend to be small. On average, the long-short return is 4.30% ($t = 5.45$) before and 1.65% ($t = 2.17$) after adjusting for ROLL, again suggesting nonnegligible profits to arbitrage by institutions.

Finally, we consider the impact of the delisting bias. Shumway (1997) shows that correct delisting returns for stocks delisted for negative reasons are often unavailable on CRSP, causing an upward bias in computed returns. This can be important since our long position loads on low-price stocks that are most likely to be delisted. In our sample, the percentage of firms in the highest LOTT quintile that are delisted in January is below 0.3%. The implied impact on quintile returns is modest. Even if

[25] For around two-thirds of our stocks we obtain a negative autocovariance, leading to a positive implied bid–ask spread. For the remaining stocks that have positive autocovariances (which imply negative bid–ask spreads), we assume the bid–ask spread equals the mean ROLL of the quintile for a given month. This adjustment tends to overestimate the impact of the bid–ask spread and, hence, yields a conservative measure of our trading profits. In unreported comparison between the ROLL measures and the relative bid–ask spread measures computed using the ISSM and The Trade and Quote (TAQ) data that are unavailable before 1983,we find that ROLL tends to overestimate the bid–ask spread by 20 to 100 percent, particularly among low-price stocks. In other words, using ROLL is likely to underestimate the returns that are net of the actual bid–ask spread.

Table VI. Daily retail and institutional order imbalance

Panel A reports the average daily retail order imbalance for the period 1983–2000 for NYSE and AMEX common stocks across three windows: early January (the first 9 days of January), late December (the last 9 days of December), and the other days of the year (mid-January to mid-December). It shows that retail investors are more bullish on stocks with a high lottery index in early January than in other days, and they rebalance portfolios by selling more lottery-type stocks in late December and buying them in early January. Retail trades are defined as trades below $5,000 using the 1991 dollar. The retail order imbalance (RIMB) is defined as the dollar amount of buyer- minus seller-initiated retail trades over the total dollar amount of buyer- and seller-initiated retail trades or $\frac{\text{buy}-\text{sell}}{\text{buy}+\text{sell}} \times 100$. Panel B reports the average daily institutional order imbalance for the period 1983–2000 for NYSE and AMEX common stocks. The institutional order imbalance (IIMB) is defined similar to RIMB but over institutional trades. Institutional trades are defined as trades above $20,000 using 1991 dollars. This panel shows that institutional investors' sentiment for lottery-type stocks does not differ between early January and other times of the year, and they do not rebalance portfolios at the turn of the year.

		Early January			Late December		
LOTT rank	Other Mean	Mean	January − other	*t* Statistic	Mean	December − other	*t* Statistic
			Panel A: retail buy order imbalance				
			Value weighted				
L (low)	0.52	−3.68	−4.20	−9.12	−4.56	−5.08	−2.83
2	0.40	−0.63	−1.02	−1.47	−5.91	−6.31	−4.81
3	−0.09	1.71	1.80	1.68	−8.22	−8.12	−8.25
4	−2.35	1.77	4.11	3.76	−14.91	−12.56	−7.86
H (high)	−5.20	3.03	8.23	5.86	−20.82	−15.63	−9.90
$H-L$	−5.72	6.71	12.79	9.15	−16.27	−9.97	−4.69
$t(H-L)$	−5.39	3.60			−13.07		
			Equal weighted				
L (low)	−1.17	−3.87	−2.52	−4.83	−6.62	−5.45	−3.80
2	−1.26	−2.55	−1.03	−1.13	−6.89	−5.63	−4.31
3	−1.95	−0.28	1.75	1.48	−9.15	−7.20	−4.71
4	−3.58	0.09	3.51	2.46	−14.77	−11.19	−7.71
H (high)	−7.23	0.21	7.15	4.36	−22.91	−15.68	−16.36
$H-L$	−6.06	4.08	10.44	7.03	−16.29	−9.41	−4.80
$t(H-L)$	−6.86	1.84			−7.94		
			Panel B: institutional buy order imbalance				
			Value weighted				
L (low)	7.08	7.29	0.21	0.30	8.15	1.08	0.98
2	4.32	6.20	1.88	1.71	6.18	1.86	3.30
3	2.37	4.48	2.11	1.81	3.91	1.54	3.20
4	0.56	2.72	2.16	1.73	2.70	2.14	2.15
H (high)	−3.57	−0.31	3.26	1.89	0.21	3.78	4.34
$H-L$	−10.64	−7.59	3.22	2.22	−7.94	3.53	3.30
$t(H-L)$	−9.05	−5.39			−5.54		

(*Continued*)

Table VI. (Continued)

LOTT rank	Other Mean	Early January Mean	Early January January − other	Early January t Statistic	Late December Mean	Late December December − other	Late December t Statistic
			Equal weighted				
L (low)	4.81	4.88	0.15	0.19	5.98	1.17	2.39
2	2.21	4.27	2.05	1.50	3.83	1.63	3.01
3	0.41	2.31	1.85	1.50	1.10	0.68	0.62
4	−1.05	2.72	3.15	2.00	0.78	1.82	1.18
H (high)	−4.98	−0.30	3.38	1.72	−0.70	4.28	5.41
H − L	−9.79	−5.17	4.72	2.23	−6.68	3.85	3.18
$t(H-L)$	−26.60	−2.31			−6.54		

we replace the missing delisting returns with −100%, the reduction in returns of the highest LOTT quintile is merely 0.3%.

4.2.c. Retail versus institutional sentiment

Next, we seek more direct evidence that individual investors are driving the January price run-ups of lottery-type stocks. Following Barber, Odean, and Zhu (2009), we study whether retail investors favor lottery-type stocks, particularly in January, by comparing the order imbalance, defined as $\frac{\text{buy}-\text{sell}}{\text{buy}+\text{sell}}$, for lottery- and nonlottery-type stocks. Our Hypothesis 3a predicts that retail orders should tilt toward buying lottery-type stocks more in January than other months. Another purpose of this test is to see if in our sample period, retail investors systematically rebalance their portfolios by selling at the year-end and buying early in the New Year, as shown by prior literature (e.g., Ritter, 1988; Starks, Yong, and Zheng, 2006) under slightly different contexts. Such evidence can further affirm our argument that portfolio rebalancing combined with a preference to gamble contributes to the high returns of lottery-type assets in the New Year.

Following Ritter (1988), we look at investors' trading behavior across three windows: early in January (the first nine trading days), late December (the last nine trading days), and the remaining days of the year from mid-January to mid-December (other days). We report in Panel A of Table VI, the retail order imbalance (RIMB) for the LOTT quintiles of stocks across the three windows using both value- and equal-weighting schemes. First, compared to the lowest LOTT quintile, RIMB is significantly lower for the highest quintile in other days by 5.72 and 6.06%, suggesting that retail investors are not bullish on lottery stocks as much as on nonlottery stocks during these days. The weaker sentiment on lottery-type stocks is even stronger in late December, possibly caused by tax-loss selling since these stocks tend to underperform over the course of the year. In early

January, however, such retail sentiment reverses; RIMB is higher for the highest LOTT quintile than for the lowest quintile by 4.08 and 6.71% with t statistics of 1.84 and 3.60. This evidence suggests that retail investors are (marginally) significantly more bullish on lottery-type stocks in early January. This corroborates our earlier evidence that, in January, lottery-type stocks earn significantly higher returns; the unusual buying behavior of retail investors can drive up prices of lottery-type stocks in January.

A possible concern is that institutions, on the other hand, can also exhibit bullish sentiment in January due to reasons other than a gambling preference, such as a risk-shifting tendency. In this case, we would not be able distinguish whether it is retail or institutional sentiment that drives January prices of lottery-type stocks. Thus, we conduct the same exercises with respect to institutional trades. Panel B of Table VI reports the institutional order imbalance (IIMB) for the LOTT quintiles across the same three windows. The picture is clear; unlike retail orders, institutional orders exhibit no January seasonality. Consistent across early January, late December, and other days, institutions are more bullish on nonlottery than lottery stocks. Thus, institutions alone are unlikely to be responsible for the abnormal returns of lottery-type stocks in January.

In addition, evidence in Panel A provides results consistent with the notion that retail investors tend to rebalance their portfolios, especially their lottery-type holdings, at the turn of the year. Panel A shows that there are significantly negative differences in the order imbalances between late December and other days, suggesting that retail investors sell all types of holdings toward the year-end. In early January, however, retail investors exhibit abnormal buying activities in the highest three LOTT quintiles but not in the lowest two LOTT quintiles, suggesting that the portfolios of retail investors are rebalanced toward lottery-type holdings. In contrast, such portfolio rebalancing is not evident for institutional investors. Relative to the order imbalance in other days, there is some evidence that institutions tend to buy more stocks in early January, but there is no evidence that they sell more in late December.[26] Again, the evidence suggests that the January outperformance of lottery-type stocks may be associated with the portfolio rebalancing behavior toward lottery-type stocks in early January of retail investors but not by that of institutional investors.

4.2.d. Multivariate regressions

So far, our evidence is consistent with the hypothesis that investor's gambling preference impacts returns of lottery-type stocks in January. Next, we use Fama–

[26] The order imbalance of institutional orders shows bearish sentiment for lottery-type stocks and bullish sentiment for non-lottery stocks in early January.

Table VII. Fama–MacBeth regressions at the firm level for January returns

Panel A shows that the three lottery feature variables and the lottery index independently and incrementally, relative to a number of well-known return predictors and measures of systematic risk, forecast individual stock returns in January. It reports the firm-level Fama and MacBeth (1973) regression results during January 1964 to January 2007 for January returns. The dependent variable is the percentage January returns of individual stocks. LOTT, LOGPRC, IVOL, EISKEW, LOGME, LOGBM, RET(−1), RET(−12, −2), and RET(−36, −13) are defined in Tables I and II. β_{MKT}, β_{SMB}, and β_{HML} refer to loadings on the three Fama–French factors, which are estimated using at least fifteen daily returns within the current January. Panel B includes the retail order imbalance (RIMB), as defined in Table V and measured in month *t*, and its interaction with the four measures of lottery stock characteristics. Panel C includes the RTP, as defined in Table II and measured in month *t*, and its interaction with the four measures of lottery stock characteristics. The control variables include all other characteristics and factor loadings in Panel A. Panel D splits the full sample based on past returns or institutional trading in the first quarter of the year and shows that the forecast power of LOTT is significant for all subsamples. Specifications (1)–(4) in Panel A are run in Panel D. WINNER refers to stocks with positive past 12-month returns as of the most recent December. LOSER refers to those with negative cumulative returns over the same horizon. IO-NBUY refers to stocks that are either untraded or net sold by institutions, defined as no net change or a net decrease in institutional ownership, in the first quarter of the current year. IO-BUY refers to stocks that are net bought by institutions, defined as a net increase in institutional ownership over the same time horizon. LOTT' is a redefined lottery feature index using only stocks in each of the subsamples. The control variables include all other characteristics and factor loadings in Panel A. All returns are expressed on a monthly basis. The R^2 is adjusted for degrees of freedom and their time series means are reported.

Panel A: regression of January returns on lottery characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LOTT	0.812 (5.92)					0.312 (3.54)				
LOGPRC		−4.441 (−6.95)					−3.244 (−5.63)			−1.774 (−2.84)
IVOL			1.462 (7.84)					0.752 (6.71)		0.300 (2.39)
EISKEW				14.582 (7.00)					13.908 (5.45)	5.689 (1.94)
LOGME					−1.554 (−7.83)	−0.930 (−7.63)	−0.164 (−1.07)	−1.065 (−6.33)	0.402 (1.42)	0.131 (0.42)
LOGBM					0.028 (0.09)	0.184 (0.63)	0.243 (0.88)	0.264 (0.93)	0.194 (0.83)	0.410 (1.19)
RET(−1)					−0.142 (−7.00)	−0.144 (−7.08)	−0.124 (−6.89)	−0.161 (−7.79)	−0.121 (−5.45)	−0.140 (−4.95)
RET(−12, −2)					−0.202 (−3.56)	−0.174 (−3.51)	−0.094 (−2.27)	−0.166 (−3.28)	−0.046 (−0.89)	−0.061 (−1.17)
RET(−36, −13)					−0.316 (−4.03)	−0.283 (−4.12)	−0.187 (−3.49)	−0.265 (−3.88)	−0.221 (−3.82)	−0.182 (−3.57)
β_{MKT}					1.135 (5.46)	1.073 (5.17)	1.029 (5.14)	1.033 (5.45)	0.934 (5.78)	0.888 (5.73)
β_{SMB}					0.283 (2.03)	0.260 (2.03)	0.265 (2.11)	0.269 (2.16)	0.288 (2.28)	0.293 (2.39)
β_{HML}					0.005 (0.03)	−0.283 (−4.12)	−0.014 (−0.07)	0.007 (−0.04)	−0.059 (−0.31)	−0.079 (−0.43)

(*Continued*)

Table VII. (Continued)

Panel A: regression of January returns on lottery characteristics

Intercept	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (%)	5	8	4	9	13	13	14	14	15	16

Panel B: interaction with retail net order imbalance (RIMB)

	(1)		(2)		(3)		(4)
LOTT	0.268 (2.66)	LOGPRC	−3.099 (−4.08)	IVOL	0.522 (3.82)	EISKEW	9.276 (3.79)
RIMB	−5.397 (−3.40)	RIMB	24.831 (6.43)	RIMB	1.709 (2.44)	RIMB	21.072 (5.26)
LOTT × RIMB	1.653 (5.31)	LOGPRC × RIMB	−7.296 (−5.95)	IVOL × RIMB	2.117 (3.97)	EISKEW × RIMB	23.494 (6.44)
Controls	Yes		Yes		Yes		Yes
Intercept	Yes		Yes		Yes		Yes
R^2 (%)	13		15		13		13

Panel C: interaction with RTP

	(1)		(2)		(3)		(4)
LOTT	0.031 (0.31)	LOGPRC	−1.724 (−3.11)	IVOL	0.308 (2.15)	EISKEW	3.109 (1.09)
RTP	−37.713 (−6.36)	RTP	3.385 (1.92)	RTP	−5.915 (−3.66)	RTP	3.322 (2.58)
LOTT × RTP	2.597 (5.94)	LOGPRC × RTP	−9.175 (−6.13)	IVOL × RTP	0.978 (4.32)	EISKEW × RTP	36.247 (6.32)
Controls	Yes		Yes		Yes		Yes
Intercept	Yes		Yes		Yes		Yes
R^2 (%)	12		13		11		12

Panel D: subsamples-based past returns and institutional trading

	WINNER	LOSER	IO-NBUY	IO-BUY
(1) LOTT’	0.350 (5.13)	0.503 (5.15)	0.458 (3.35)	0.479 (3.66)
(2) LOGPRC	−2.036 (5.81)	−3.700 (5.57)	−3.116 (3.57)	−3.505 (3.82)
(3) IVOL	0.416 (3.83)	0.606 (3.97)	0.634 (4.24)	1.137 (3.26)
(4) EISKEW	8.271 (4.52)	15.848 (4.79)	13.783 (4.79)	14.632 (3.79)
Specifications (1)–(4)				
Controls	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes

MacBeth regressions to study this effect, which allows us to control for other variables that are known to forecast January returns.

In the Fama–MacBeth regressions, we control for stock price, size and book-to-market equity, and past short- and long-run returns. The long-term reversal effect is most pronounced (Conrad and Kaul, 1993; Loughran and Ritter, 1996) and the short-run momentum effect is reversed (Jegadeesh and Titman, 1993, 2001) in January. We also see if IVOL and EISKEW have incremental power to forecast January returns after considering the roles of price and size since low-price stocks outperform high-price stocks in January and largely drive the small-firm-in-January effect (Keim, 1983; Kross, 1985; Bhardwaj and Brooks, 1992).

In Panel A of Table VII, we report the results of firm-level Fama–MacBeth regressions. The dependent variable is January stock returns. The independent variables are LOTT, LOGPRC, IVOL, EISKEW, LOGME, LOGBM, RET(−1), RET(−12, −2), RET(−36, −13), and the loadings on the three Fama–French factors: the market beta (β_{MKT}), SMB loading (β_{SMB}), and HML loading (β_{HML}). The factor loadings are estimated from the three-factor model over at least 15 days in the current January and control for systematic risk (excluding the loadings does not change the results). We include each of the lottery feature measures alone and with these controls to examine whether the lottery-stock-in-January effect is distinct from known January effects.

The Fama–MacBeth regression results, reported in Panel A of Table VII, give strong support to our hypothesis. They show that the lottery feature measures forecast firm-level January returns alone and after including control variables. When used alone, all four lottery feature measures significantly forecast returns in the expected direction, negative for LOGPRC and positive for IVOL, EISKEW, and LOTT. The coefficient on LOTT is 0.812 ($t = 5.92$). Considering that LOTT ranges from 0 to 19, a change from the lowest to the highest ranking increases the mean January stock return by more than 16%. Each of the four lottery feature measures remains statistically significant when we include control variables. The coefficient on LOTT, 0.312 ($t = 3.54$), is reduced by more than 60% but remains highly significant; the change from the bottom to the top LOTT ranking corresponds to a marginal effect of more than 6% per January. Therefore, the outperformance of lottery-type stocks in January is incremental to previously known January seasonality.

The results also show that the lottery-stock-in-January effect is more than a January price effect. IVOL and EISKEW both remain positive and statistically significant after adding them to regressions with LOGPRC and other controls. That is, both volatility and skewness play an indispensable role in selecting lottery-type stocks. More interestingly, comparing the base Specification (5) with no lottery feature variables to those with them, the coefficient on log firm size is reduced in magnitude and in some cases even reverses signs. Specifically, including LOTT, LOGPRC, and IVOL individually reduces the coefficient of LOGME from −1.554

to −0.930, −0.164, and −1.065, respectively. Including EISKEW or the three lottery features reverses the LOGME coefficient to an insignificant 0.402 and 0.130, respectively. Since returns have a significant relation with EISKEW and the inclusion of EISKEW eliminates the negative relation between returns and size, skewness may be causing the January size effect in security returns.

4.2.e. Retail trading, sentiment, and stock returns

Next, we test Hypothesis 3b that the January outperformance of lottery-type stocks should be stronger among those more bullishly thought of or more actively traded by retail investors. Thus, we include the retail order imbalance (RIMB) and its interaction with the LOTT index (or the components of the index) as additional independent variables in the Fama–MacBeth regressions. In a different specification, we also include the RTP and its interaction with the LOTT index (or its components) in the Fama–MacBeth regressions. We expect the interaction term to be positive for LOTT, IVOL, and ESKEW but negative for LOGPRC.

The results are reported in Panels B and C of Table VII. Consistent with our hypothesis, controlling for the previous firm characteristics, we find that the coefficients on the interactions between RIMB and the lottery features have the expected sign and are highly significant, while the lottery feature variables remain significant. This suggests that lottery-type stocks that are bullishly thought of and thus bought by retail investors earn even higher returns in January than the already high returns earned by general lottery-type stocks. Similar results are obtained by using RTP. In Panel C, the interaction terms are all significant, while two out of four lottery feature variables become insignificant. This result further confirms that retail trading activity is essential in driving lottery-type stock returns in January.

4.2.f. Tax-loss selling, institutional window dressing, and risk shifting

To distinguish our gambling-based hypothesis from traditional hypotheses for the January effect of lottery stocks, we examine whether the lottery-stock-in-January effect is driven by tax-loss selling or window dressing, both of which predict that high returns in January should exclusively occur to past loser stocks. We separately run Fama–MacBeth regressions for past winner and loser stocks and report the results in Panel D of Table VII, where we define winners (losers) as stocks with positive (negative) 12-month cumulative returns as of the end of December. For the two separate groups, we add each of the lottery feature variables to the set of controls used in Panel A of Table VII. All lottery feature variables remain significant in both winner and loser groups, suggesting that the lottery-stock-in-January effect occurs despite past returns. Thus, tax-loss selling and window dressing do not fully explain this effect. This effect, however, is stronger for losers; the coefficient on LOTT is 0.503 ($t = 5.15$) for losers and 0.350 ($t = 5.13$) for winners.

Next, we test the risk-shifting hypothesis. Following Ng and Wang (2004), we separate stocks into two groups based on the change in institutional ownership over the first quarter of each year. We characterize each stock as a net buy if institutional holdings increase and no-net-buy otherwise. We then run a Fama–MacBeth regression on each of the lottery feature variables together with the controls. Panel D of Table VII shows significant coefficients on the lottery feature variables: 0.479 ($t =$ 3.66) among the net-buy group and 0.458 ($t = 3.35$) among the no-net-buy group. In other words, the lottery-stock-in-January effect persists even when institutions are not buying, again suggesting that individual gambling preference makes a price impact. In unreported tests, we do not find that stocks with the strongest lottery features tend to announce more favorable earnings news. This is consistent with Peterson (1990), who suggests that information revelation cannot explain the broad January phenomena in stock returns.

4.2.g. Daily return and retail sentiment surrounding New Year's Day

So far, we have focused on monthly returns. We proceed to test Hypothesis 3c by studying daily returns and retail sentiment surrounding New Year's Day.

In Panel A of Figure 3, we plot the equal-weighted average daily returns of the highest LOTT quintiles over a 40-trading-day window surrounding January 1. The highest LOTT quintile experiences a small price run-up over three trading days around Christmas, with a daily appreciation as high as 100 basis points. Later, it has a large price run-up over a 5-day window from 1 day before through 4 days after New Year's Day, with a daily appreciation as high as 300 basis points in the first trading day of the New Year. The abnormal return gradually recedes until the end of January. The return effect is at least partly caused by retail investors. The retail-buy lottery-type stocks, defined as the top LOTT quintile of stocks with positive order imbalance (RIMB) in January, and the institution-sell lottery stocks, defined as the top LOTT quintile of stocks that have decreased institutional ownership in the first quarter of the current year, depict a very similar picture as the general lottery-type stocks. By contrast, the nonlottery stocks in the lowest LOTT quintile have returns indistinguishable from zero at the turn of the year.[27]

Panel B of Figure 3 provides further evidence that retail investor sentiment contributes to the turn-of-the-year effect. Here, we plot the daily retail buy-minus-sell turnover of the top and bottom LOTT quintiles of stocks. Retail buy-minus-sell

[27] In unreported analyses we find the return behavior of post winners stocks in the top LOTT quintile is similar to the general lottery-type stocks, suggesting that this return effect is not solely caused by past loser stocks.

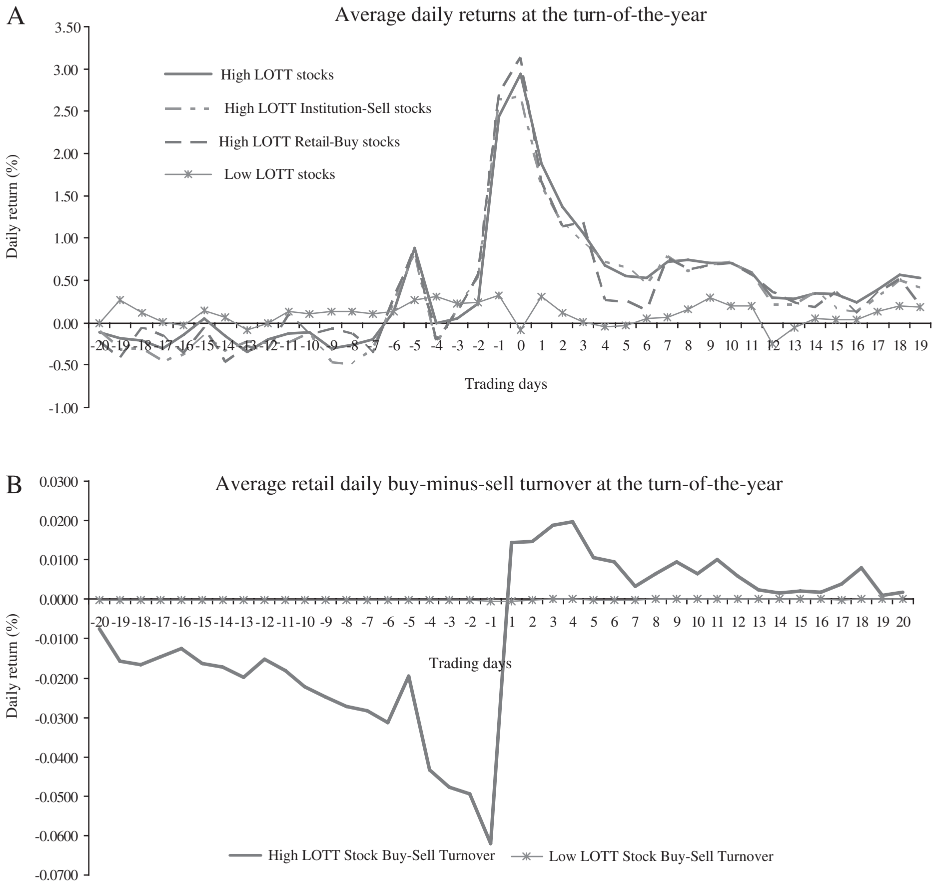

Figure 3. Daily returns and excess turnover of lottery-type stocks at the turn-of-the-year. Panel A depicts average daily percentage returns for 20 trading-days before through 20 trading-days after January 1. Day 0 is the first trading-day in January. Stocks must have positive daily volume to be included. High LOTT stocks are defined as the top quintile of stocks based on the lottery-feature index (LOTT), as in Table II. High LOTT Institution- Sell stocks are lottery-type stocks (defined as above) without a net increase in institutional ownership during the first quarter of the current year. High LOTT Retail-Buy stocks are lottery-type stocks with a positive monthly order imbalance in January, as defined in Table VI. Low LOTT stocks are defined as the bottom quintile based on LOTT. Panel B depicts the abnormal buy-minus-sell retail percentage turnover for 20 trading-days before through 20 trading-days after January 1. The buy-minus-sell retail turnover is defined as the difference between the total buy dollar volume and the total sell dollar volume of retail investors (with per trade transaction below $5000 in 1991 dollar), scaled by market cap at the end of the day. The abnormal buy-minus-sell retail turnover is the daily buy-minussell retail turnover during the event window minus the average daily buy-minus-sell retail turnover in non-event days of the previous year from February to November. The portfolio composition remains constant across the trading days. Daily returns and abnormal buyminus- sell retail turnover are equal-weighted. Plots involving institutional ownership cover the period 1980–2007. Those involving the retail order data cover the period 1983–2000. Others cover the period 1963–2007. Only common stocks traded on the NYSE and AMEX are included in the analyses.

turnover is the dollar amount of buyer- minus seller-initiated trades over the market cap at the end of the day. It is similar to the order imbalance but accounts for the fact that retail trades are only a fraction of the total trades.[28] As shown in Panel B, the retail buy-minus-sell turnover changes from negative to positive at the turn of the New Year (from day $t-1$ to $t+1$) for both the top and bottom LOTT quintiles. The sentiment shift is significantly stronger for lottery than nonlottery stocks. Specifically, the daily buy-minus-sell turnover switches from -0.06 to 0.01% in one day for lottery stocks but the change is close to zero for nonlottery stocks (because retail trades are only a tiny fraction of total trades for these stocks). The evidence depicts a reversal from bearish to bullish sentiment for retail investors at the turn of the year. This coincides with the price surge in lottery stocks over several days after New Year's Day.

Although our anecdotal evidence suggests that gambling mentality is likely strong from Christmas through early in the New Year, the retail sentiment revealed in the order imbalance does not turn bullish until after New Year's Day. This excess selling behavior toward the year-end is likely caused by the tax-loss selling in lottery-type stocks since these stocks tend to underperform over the course of the year. This, however, does not preclude the possibility that investors' gambling preference guides new money into lottery-type assets after Christmas and before New Year's Day.[29]

4.2.h. Individuals versus institutions

Next, we study whether lottery-type stocks purchased by individuals have worse returns in the long run than those bought by institutions. We track separately the top and bottom LOTT quintiles, formed at the end of the December of year $s-1$, based on whether retail (institutional) investors are net buyers, either according to the ISSM/TAQ data in January or the 13f data in the first quarter of year s. We compute the cumulative equal-weighted quintile returns from February to December of year s by keeping the composition of each quintile constant. Individual stock returns incorporate delisting returns and, in the case of a missing delisting return, we use -30% (Shumway, 1997). The purpose is to measure the expected returns

[28] For example, one may observe similar buy-sell-imbalance for lottery and non-lottery stocks while the retail trades account for 80% of volume among lottery stocks but 20% among non-lottery stocks. In this case, using the buy-sell-imbalance figures will not be as informative as the buy-sell-turnover to understand the price impact of the retail trades. In contrast, the buy-sell turnover accounts for not only the buy-sell-imbalance but also the retail trades as the fraction of total trades since it is the product of the two.

[29] The fact that lottery-type stocks experience some price runup toward the year-end may be caused by the buying behaviors of institutions. As shown in Panel B of Table VI, institutions tend to buy, not sell, stocks, including lottery-type stocks in late December. There is much anecdotal evidence that institutions take advantages of the January effect by buying towards the year-end. See, for example, "January effect under scrutiny," by Reena Aggarwal and Pietra Rivoli, *Portfolio Management*, p.25, May 30, 1988.

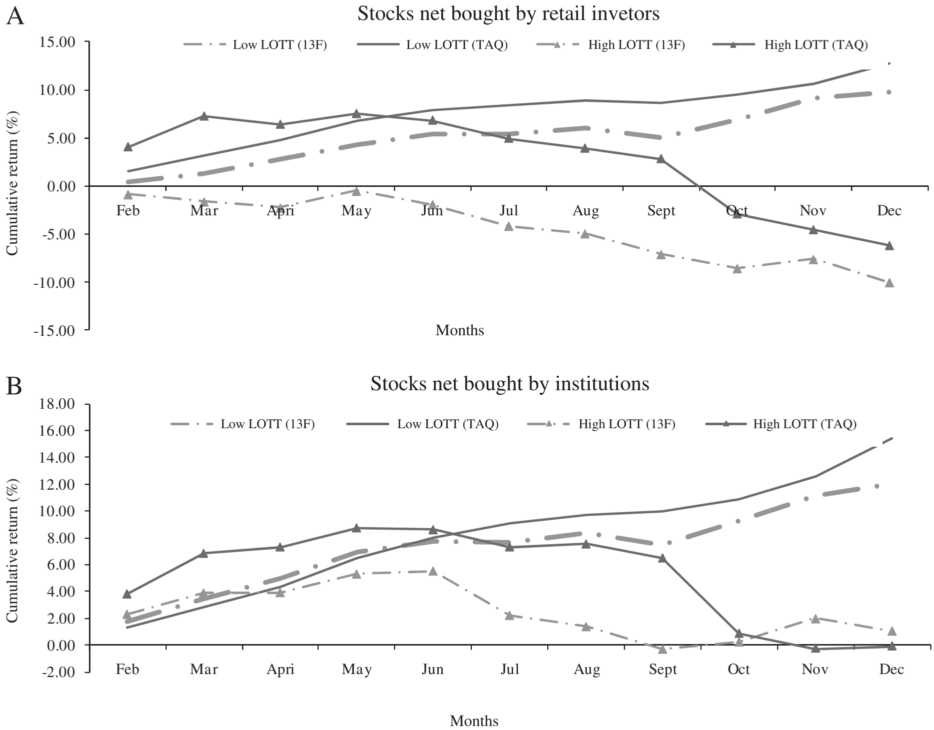

Figure 4. Cumulative buy-and-hold returns from February to December based on the lottery index and changes in institutional ownership. The figure in Panel A depicts the cumulative percentage value-weighted returns on the highest and the lowest lottery index (LOTT) quintiles of stocks for which retail investors are net buyers. Panel B depicts these returns for stocks for which institutional investors are net buyers. Purchases are from 13F data in the first quarter of the year (1981-2007) or the ISSM/TAQ data in January of the year (1989-2000). The LOTT quintiles are formed at the end of the December preceding January of year s. We keep the composition of each quintile constant and compute value-weighted quintile monthly returns from February to December of year s. Then we cumulate these mean quintile monthly returns from February through December. Delisting returns are incorporated in the month of delisting. If delisting returns are missing they are replaced by −30% (Shumway 1997). The figures show that for the year excluding January, stocks in the highest LOTT quintile are expected to have lower returns than those in the lowest LOTT quintile. In particular, among stocks for which individual investors are net buyers, the lottery-type stocks are expected to deliver negative returns by the end of the year.

of lottery-type stocks that are identified at the turn of the year using the realized returns in the rest of the year after January.

Figure 4 depicts an interesting pattern. Consistent with the findings in Table V, the highest LOTT quintile substantially underperforms the lowest quintile from February through December regardless whether we identify the net bought stocks by individuals or institutions using either the ISSM/TAQ or the 13f data. Also as predicted, the underperformance of lottery-type stocks is stronger for those purchased by individuals than those purchased by institutions. In Panel A, stocks that are net bought by retail investors have significant negative cumulative returns (−6.22 to −9.98%) as of the end of December. In contrast, in Panel B, those net bought by institutions have cumulative returns indistinguishable from zero (−0.09 to −1.07%). Both groups of lottery-type stocks underperform the comparable nonlottery-type groups. Like playing lotteries, investors, particularly individual investors, expect to earn low or negative returns. However, individual investors are significantly worse off through lottery-stock picking. The evidence confirms our hypothesis that individual investors gamble despite negative expected returns.

4.3 CHINESE STOCK RETURNS AROUND THE CHINESE NEW YEAR

Finally, we examine Chinese stocks around the Chinese New Year. Chinese celebrate the Chinese New Year more seriously than January 1 and have a tradition to gamble in the New Year. Thus, we hypothesize that Chinese are more likely to express their gambling preference at the turn of the Chinese New Year rather than January 1.

4.3.a. Market performance

To test Hypothesis 4a, we report the mean monthly returns on the equal-weighted market portfolio in Panel A of Table VIII for all months, the Chinese New Year's (CNY) Month, January (JAN), and all other months from March through December. From 1994 to 2006, the average market return is highest during the Chinese New Year's Month with a mean monthly return of 5.92% and lowest during January, −1.53%, which excludes the Chinese New Year's Month.[30] The mean returns are mildly positive for the remaining months; the equal-weighted Chinese stock market portfolio does not exhibit a January effect, but it does have a Chinese New Year effect.[31] This is consistent with Girardin and Liu (2005) and Hsu (2005), who find that Chinese stock markets exhibit a weak January effect. The result confirms the hypothesis based on the New Year's gambling mentality that the Chinese New Year affects stock returns more than the January 1 New Year.

4.3.b. Individual stock performance

Next, we test Hypothesis 4b by examining the returns of Chinese lottery-type stocks. Due to the rather short time period (1994–2006), we use firm-level pooled

[30] The two major Chinese exchanges, Shanghai and Shenzhen, were established at the end of 1990. DataStream starts reporting data in 1992. Our analyses start in 1994 to ensure a sufficiently large cross section of stocks. From 1992 through 1993, there are few stocks reported in DataStream (less than 30 by the end of 1992 and less than 100 by the end of 1993). The results are stronger if we include data from 1993.

[31] In unreported tests, we find that the value-weighted Shanghai and Shenzhen indexes exhibit a weak January effect but a strong Chinese New Year effect.

Table VIII. Chinese stock returns: January versus the Chinese New Year

Panel A reports the average monthly percentage returns of the equal-weighted market portfolio over three time intervals for the period January 1994 to December 2006 for Chinese A shares. The time periods are the Chinese New Year's Month (CNY), defined as a 22-trading-day window beginning the first trading-day of the Chinese New Year, trading days in January that precede the New Year's Month (JAN), and other times of the year from March through December (OTHER). Other times includes all trading-days in April through December, and trading-days in February and March after the New Year's Month. All such February returns are treated as if they occur in March. To account for the variable number of trading days in January and March, portfolio returns are divided by the number of valid trading days within the month and multiplied by 22. Thus, all returns are expressed on a monthly basis. Panel B reports the results of a pooled cross-sectional regression at the firm level. The dependent variable is the monthly individual stock returns minus the equal-weighted monthly market portfolio return. The two lottery feature variables are logarithmic stock price, LOGPRC, and idiosyncratic volatility, IVOL. PRC is defined as the closing price of the last trading day of the previous month. IVOL is the standard deviation of at least fifteen daily residual returns in the preceding month from the regression of daily returns on both the Shanghai and Shenzhen A share indices. When the number of trading days in the preceding month is less than 15 for the overall market, IVOL from month $t-2$ is used. LOGPRC $\times$ CNY is equal to LOGPRC when the return is measured over the Chinese New Year's Month and 0 otherwise. LOGPRC $\times$ JAN is equal to LOGPRC when the return is measured over January and 0 otherwise. Similar definitions hold for IVOL $\times$ CNY and IVOL $\times$ AN. *t* Statistics reported in parentheses are based on standard errors that cluster over both firm and time.

Panel A: equal-weighted market returns

ALL MONTHS	CNY	JAN	OTHER
1.18 (0.94)	5.92 (2.64)	−1.53 (−0.36)	0.97 (0.68)

Panel B: pooled cross-sectional regression

	(1)	(2)	(3)
LOGPRC	−0.889 (−5.78)		−0.860 (−5.55)
LOGPRC × CNY	−1.283 (−2.50)		−1.344 (−2.57)
LOGPRC × JAN	−0.466 (−0.08)		−0.496 (−0.90)
IVOL		−0.479 (−3.86)	−0.419 (−3.34)
IVOL × CNY		0.988 (1.80)	1.039 (2.05)
IVOL × JAN		0.495 (1.09)	0.456 (1.06)
Intercept	Yes	Yes	Yes
R^2 (%)	1.10	0.19	1.26

cross-sectional regressions instead of portfolio sorts to examine the return seasonality of lottery-type stocks. Panel B of Table VIII reports the results. The dependent variable is the monthly stock return minus the equal-weighted monthly market portfolio return. The independent variables include LOGPRC, IVOL, and their interactions with a Chinese New Year's Month dummy and January dummy. The interaction term, LOGPRC $\times$ CNY, equals LOGPRC if the return is in the Chinese New Year's Month and zero otherwise. Similar definition applies to IVOL $\times$ CNY, LOGPRC × JAN, and IVOL × JAN. We run three specifications. In one, we include three variables associated with LOGPRC. In the second, we include three variables associated with IVOL. In the third, we include all six. The t statistics are based on standard errors that cluster over time and firm dimensions.

In Panel B of Table VIII, we find strong evidence that lottery-type Chinese stocks outperform in the Chinese New Year's Month but not January. The coefficient on LOGPRC × CNY is always negative (−1.283 and −1.344) and significant at the 1% level. The coefficient on IVOL × CNY is always positive (0.988 and 1.039) and significant at the 10% and the 5% levels. In contrast, none of the interaction terms with the January dummy are significant. Our results are consistent with the hypothesis that Chinese investors prefer lottery-type stocks at the start of the New Year and not necessarily January unless the two periods coincide.

5. Wealth Transfer and the Survival of Gamblers in Financial Markets

If the gamblers in financial markets are predominately retail investors, as suggested by our evidence, then the New Year effect we document implies a substantial wealth transfer from unsophisticated retail investors to sophisticated and, maybe, institutional ones. Using 2007 as an example, the highest LOTT quintile has a total market cap of $98 billion (less than 1% of our total market cap). Using the 2007 January value-weighted abnormal $H - L$ return of 2.95% and the turnover for the highest LOTT quintile of 8.68%, the wealth transfer is about $251 million ($98 billion × 2.95% × 8.68%). In the option markets, the wealth transfer is more. Again for 2007, buyers of the OTM call options have a net payout over $6 billion to call writers, with about $900 million in January alone.[32] However, such net payouts for gaming purposes may be smaller since they include premiums paid for hedging purposes. A better measure of wealth transfer due to gambling preference is from direct gaming markets, where the magnitude is even greater. In 2007, the total US casino gaming revenue is $34.13 billion and the total gaming revenue including casinos, sports, horse racetracks, etc. is $92.3 billion.[33] In other words, Americans (and others) spend substantially to satiate gambling preferences. The direct wealth

[32] The $6 billion estimate is based on the total option volume in 2007 and parameter estimates of Pan and Poteshman (2006). We first multiply the total option volume reported by the OCC by 15% to obtain the estimated volume of OTM call contracts. We then multiply the number of OTM call contracts by 0.9 (the percentage of time the option expires worthless), and then $82 (the average premium paid per contract). Then we subtract a factor of 0.3 off this total premium to account for the profits on the options that expire in the money (10% of contracts expire in the money and pay off 300% of the premium on average) (Doran and Fodor 2009).

[33] See the website of American Gaming Association, http://www.americangaming.org, for gaming industry facts and statistics.

transfer through gaming is even more substantial than the indirect wealth transfer through purchasing lottery-like assets in financial markets.

Given the large magnitude of wealth transfer in financial markets caused by gambling preferences, a possible question about our New Year's results is why gamblers can survive for decades after consistently large wealth transfers from them to sophisticated investors. In general, this pertains to the question of why financial anomalies can persist in a competitive market. Several behavioral explanations address such criticism.

First, security trading is not a game between a sophisticated investor and an unsophisticated one but between groups of investors that continuously enter and exit the markets with a wide spectrum of informativeness and sophistication. Since the feedback mechanisms in financial markets can be vague and delayed, particularly for low-price and high-volatility stocks and OTM options, generations of unsophisticated investors may not learn from their own experience or that of previous generations. Even when natural selection drives out irrational investors in the limit (e.g., Luo, 1998), the process can be extremely slow, taking hundreds of years, and during this process irrational investors may dominate (Yan, 2008).

Second, unsophisticated investors can inadvertently take excessive risk due to their preferences for security characteristics that are correlated with risk or their underestimation of risk because of overweighting of private information.[34] Over time, they may earn higher risk premiums that partially compensate for irrational trading losses (DeLong *et al.* 1990; Hirshleifer and Luo, 2001). Thus, even when traders' wealth constraints are a key determinant of their demand, irrational investors' portfolios may not quickly deplete. Third, investor learning need not be rational. Successful investors can learn to be irrational due to biased self-attribution (Gervais and Odean, 2001), implying a reversed wealth transfer from these once sophisticated, but later overconfident, investors. Finally, limits of arbitrage are particularly important barriers for low-price and high-volatility stocks for which institutional investors are known to be averse to (e.g., Falkenstein, 1996). Thus, arbitrage is likely incomplete for this segment of markets.

In short, there are several opposing forces that can impede or delay the process of the market converging to full efficiency in the long run. Therefore, over a period of several decades, it is unknown whether rational investors would drive out of irrational investors.

6. Conclusions

Gambling is a built-in preference of some individuals and tends to be stronger surrounding the New Year. This paper provides new evidence showing that such

[34] In our context, gamblers prefer highly volatile stocks, which are also riskier.

preference exists in financial markets and has a strong price impact on assets with lottery features in the New Year. We show that OTM call options and lottery-type stocks in the USA and China (where the New Year is not January 1) have abnormally high prices, returns, and trading volume at the turn of the New Year and retail sentiment and trading drive this New Year seasonality. This seasonality contributes to, but differs from, previously known January effects and carries important implications for trading strategies. Tax-loss selling, window dressing, and risk shifting do not explain all the gambling effect.

We show that gaming revenues from the Las Vegas Strip and lottery sales in Mega Millions and Powerball exhibit January seasonality. We also find novel seasonality in option implied volatility, volume, and behavior and sentiment of retail versus institutional investors. The underperformance of high idiosyncratic volatility (Ang *et al.*, 2006) and idiosyncratic skewness (Boyer, Mitton, and Vorkink, 2010) stocks are shown to be a pure non-January phenomenon. Furthermore, the New Year effect of Chinese stocks poses an interesting puzzle for tax-loss-selling and institutional trading–based hypotheses.

References

Ackert, L. F., Charupat, N., , Church, B. K., and Deaves, R. (2006) An experimental examination of the house money effect in a multi-period setting, *Experimental Economics* **9**, 5–16.

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

Arkes, H. R., Hirshleifer, D., Jiang, D., and Lim, S. (2008) Reference point adaptation: tests in the domain of security trading, *Organizational Behavior and Human Decision Processes* **105**, 67–81.

Bakshi, G. and Kapadia, N. (2003) Volatility risk premium embedded in individual equity options: some new insights, *Journal of Derivatives* **10**, 45–54.

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

Barber, B. M., Odean, T., and Zhu, N. (2009) Do retail trades move markets? *Review of Financial Studies* **22**, 151–186.

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

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

Barberis, N., Huang, M., and Santos, T. (2001) Prospect theory and asset prices, *Quarterly Journal of Economics* **116**, 1–53.

Barberis, N. and Xiong, W. (2008) Realization utility. Working paper, Princeton University.

Battalio, R., Hatch, B., and Jennings, R. (2004) Toward a national market system for U.S. exchange-listed equity options, *Journal of Finance* **59**, 933–962.

Benartzi, S. and Thaler, R. (1995) Myopic loss aversion and the equity premium puzzle, *Quarterly Journal of Economics* **110**, 73–92.

Bhardwaj, R. and Brooks, L. (1992) The January anomaly: effects of low share price, transactions costs, and bid-ask bias, *Journal of Finance* **47**, 553–575.

Bollen, N. and Whaley, R. (2004) Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* **59**, 711–753.
Boyer, B., Mitton, T., and Vorkink, K. (2010) Expected idiosyncratic skewness, *Review of Financial Studies* **23**, 169–202.
Brandt, M., Brav, A., Graham, J., and Kumar, A. (2010) The idiosyncratic volatility puzzle: time trend or speculative episodes? *Review of Financial Studies* **23**, 863–899.
Brown, P., Keim, D. B., Kleidon, A. W., and Marsh, T. A. (1983) Stock return seasonalities and the tax-loss selling hypothesis: analysis of the arguments and Australian evidence, *Journal of Financial Economics* **12**, 105–127.
Brunnermeier, M. K., Gollier, C., and Parker, J. A. (2007) Optimal beliefs, asset prices, and the preference for skewed returns, *American Economic Review* **97**, 159–165.
Cargill, T. F. and Eadington, W. R. (1978) Nevada's gaming revenues: time characteristics and forecasting, *Management Science* **24**, 1221–1230.
Conrad, J. and Kaul, G. (1993) Long-term market overreaction or biases in computed returns? *Journal of Finance* **48**, 39–63.
Coval, J. D. and Shumway, T. (2005) Do behavioral biases affect prices? *Journal of Finance* **60**, 1–34.
D'Mello, R., Ferris, S. P., and Hwang, C. Y. (2003) The tax-loss selling hypothesis, market liquidity, and the price pressure around the turn-of-the-year, *Journal of Financial Markets* **6**, 73–98.
DeLong, J. B., Shleifer, A., Summers, L., and Waldmann, R. J. (1990) Noise trader risk in financial markets, *Journal of Political Economy* **98**, 703–738.
Doran, J. S. and Fodor, A. (2009) Firm specific option risk and implications for asset pricing, *Journal of Risk* **12**, 17–52.
Falkenstein, E. (1996) Preferences for stock characteristics as revealed by mutual fund holdings, *Journal of Finance* **51**, 111–135.
Fama, E. F. and French, K. R. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**, 3–56.
Fama, E. F. and MacBeth, J. D. (1973) Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* **81**, 607–636.
Friedman, M. and Savage, L. J. (1948) The utility analysis of choice involving risk, *Journal of Political Economy* **56**, 279–304.
Gervais, S. and Odean, T. (2001) Learning to be overconfident, *Review of Financial Studies* **14**, 1–27.
Girardin, E. and Liu, Z. (2005) Bank credit and seasonal anomalies in China's sock markets, *China Economic Review* **16**, 465–483.
Gultekin, M. N. and Gultekin, N. B. (1983) Stock market seasonality: international evidence, *Journal of Financial Economics* **12**, 469–481.
Harvey, C. R. and Siddique, A. (2000) Conditional skewness in asset pricing tests, *Journal of Finance* **55**, 1263–1295.
Haug, M. and Hirschey, M. (2006) The January effect, *Financial Analysts Journal* **62**, 78–88.
Haugen, R. A. and Lakonishok, J. (1992) The Incredible January Effect: the Stock Market's Unsolved Mystery. 2nd edition. Irwin Professional Publishing, Willowbrook, Illinois.
Hirshleifer, D. and Luo, G. Y. (2001) On the survival of overconfident traders in a competitive security market, *Journal of Financial Markets* **4**, 73–84.
Hsu, C.-W. (2005) Is there a January effect in the greater China area? MBA thesis, Simon Fraser University.
Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance* **48**, 65–91.
Jegadeesh, N. and Titman, S. (2001) Profitability of momentum strategies: an evaluation of alternative explanations, *Journal of Finance* **56**, 699–720.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk, *Econometrica* **47**, 263–292.
Kato, K. and Schallheim, J. S. (1985) Seasonal and size anomalies in the Japanese stock market, *Journal of Financial and Quantitative Analysis* **20**, 243–260.
Keim, D. B. (1983) Size related anomalies and stock return seasonality: further evidence, *Journal of Financial Economics* **12**, 13–32.
Keim, D. B. (1989) Trading patterns, bid-ask spreads and estimated stock returns: the case of common stocks at calendar turning points, *Journal of Financial Economics* **25**, 75–97.
Keim, D. B. and Madhavan, A. (1998) The cost of institutional equity trades, *Financial Analysts Journal* 50–69.
Kross, W. (1985) The size effect is primarily a price effect, *Journal of Financial Research* **8**, 169–179.
Kumar, A. (2009) Who gambles in the stock market? *Journal of Finance* **64**, 1889–1933.
Kumar, A. and Lim, S. S. (2008) How do decision frames influence the stock investment decisions of individual investors? *Management Science* **54**, 1052–1064.
Lee, C. M. C. and Ready, M. J. (1991) Inferring trade direction from intraday data, *Journal of Finance* **46**, 733–746.
Liu, Y-J, Tasi, C.-L., Wang, M.-C., and Zhu, N. (2010) Prior consequences and subsequent risk taking: new field evidence from the Taiwan futures exchange, *Management Science* **56**, 606–620.
Loughran, T. and Ritter, J. R. (1996) Long-term market overreaction: the effect of low-priced stocks, *Journal of Finance* **51**, 1959–1970.
Luo, G. Y. (1998) Market efficiency and natural selection in a commodity futures market, *Review of Financial Studies* **11**, 647–674.
Markowitz, H. (1952) Portfolio selection, *Journal of Finance* **7**, 77–91.
Maxwell, W. F. (1998) The January effect in the corporate bond market: a systematic examination, *Financial Management* **27**, 18–30.
Mitton, T. and Vorkink, K. (2007) Equilibrium underdiversification and the preference for skewness, *Review of Financial Studies* **20**, 1255–1288.
Ng, L. and Wang, Q. (2004) Institutional trading and the turn-of-the-year effect, *Journal of Financial Economics* **74**, 343–366.
O'Connell, P. and Teo, M. (2007) Institutional investors, past performance, and dynamic loss aversion, *Journal of Financial and Quantitative Analysis* **44**, 155–188.
Pan, J. and Poteshman, A. M. (2006) The information in option volume for future stock prices, *Review of Financial Studies* **19**, 871–908.
Peterson, D. R. (1990) Stock return seasonalities and earnings information, *Journal of Financial and Quantitative Analysis* **25**, 187–201.
Ritter, J. R. (1988) The buying and selling behavior of individual investors at the turn of the year, *Journal of Finance* **43**, 701–717.
Roll, R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* **39**, 1127–1139.
Rozeff, M. S. and Kinney, W. K. (1976) Capital market seasonality: the case of stock returns, *Journal of Financial Economics* **3**, 379–402.
Schwert, G. W. (2003) Anomalies and market efficiency, in: G. Constantinides, M. Harris, R. Stulz (eds.), *Handbook of the Economics of Finance*. Chap. 17, Amsterdam, The Netherlands: Elsevier Science Ltd, 939–974.
Shefrin, H. and Statman, M. (2000) Behavioral portfolio theory, *Journal of Financial and Quantitative Analysis* **35**, 127–151.
Shumway, T. (1997) The delisting bias in CRSP data, *Journal of Finance* **52**, 327–340.

Sias, R. W. and Starks, L. T. (1997) Institutions and individuals at the turn-of-the-year, *Journal of Finance* **52**, 1543–1562.
Starks, L. T., Yong, L., and Zheng, L. (2006) Tax-loss selling and the January effect: evidence from municipal bond closed-end funds, *Journal of Finance* **61**, 3049–3067.
Statman, M. (2002) Lottery players/stock traders, *Financial Analysts Journal* **58**, 14–21.
Statman, M. (2004) The diversification puzzle, *Financial Analyst Journal* **60**, 44–53.
Thaler, R. H. (1985) Mental accounting and consumer choice, *Marketing Science* **4**, 199–214.
Thaler, R. H. and Johnson, E. J. (1990) Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice, *Management Science* **36**, 643–660.
Van den Bergh, W. M. and Wessels, R. (1985) Stock market seasonality and taxes: an examination of the tax-loss selling hypothesis, *Journal of Business Finance and Accounting* **12**, 515–530.
Weber, M. and Camerer, C. F. (1998) The disposition effect in securities trading: an experimental analysis, *Journal of Economic Behavior & Organization* **33**, 167–184.
Xi, C. (2006) Institutional shareholder activism in China: law and practice, *International Company and Commercial Law Review* **17**, 1–46.
Yan, H. (2008) Natural selection in financial markets: does it work? *Management Science* **54**, 1935–1950.
Zhang, Y. (2005) Individual skewness and the cross-section of average stock returns, Working paper, Yale University.