Attention Allocation and Return Co-Movement:

Evidence from Repeated Natural Experiments*

Shiyang Huang†

Yulin Huang‡

Tse-Chun Lin §

May 27, 2016

Abstract

We study how attention allocation affects the composition of market/industry and firm-specific information in stock prices via repeated natural experiments. Using large jackpots of Taiwanese nationwide lotteries as exogenous shocks to investors' attention, we find: (1) individual stock returns co-move more with the market/industry returns on large jackpot days; (2) large jackpots have stronger effects on stock returns' co-movement with the market than with the industry; (3) the effect of large jackpots on return co-movement is stronger for stocks preferred by retail investors; and (4) the market under-reacts to firms' earnings surprises on large jackpot days and reverts within one week. Our findings are consistent with the existing theory that attention-constrained investors focus more on market- and sector-level information than on firm-specific information.

---

* We appreciate the helpful comments from Yifei Mao, Yu Yuan, Dexin Zhou, Shaojun Zhang, and seminar participants at University of Hong Kong. Tse-Chun Lin gratefully acknowledges research support from the Faculty of Business and Economics at the University of Hong Kong and the Research Grants Council of the Hong Kong SAR government. All errors remain the responsibility of the authors.

† Faculty of Business and Economics, The University of Hong Kong. Tel.: (+852) 3917-8564. E-mail address: huangsy@hku.hk.

‡ Faculty of Business and Economics, The University of Hong Kong. Tel.: (+852) 5220-5196. E-mail address: u3002669@connect.hku.hk.

§ Faculty of Business and Economics, The University of Hong Kong. Tel.: (+852) 2857-8503. E-mail address: tsechunlin@hku.hk.

JEL Classification: G11, G12, G14

Keywords: limited attention, attention allocation, return co-movement, earnings surprises

Attention Allocation and Return Co-Movement:

Evidence from Repeated Natural Experiments

Abstract

We study how attention allocation affects the composition of market/industry and firm-specific information in stock prices via repeated natural experiments. Using large jackpots of Taiwanese nationwide lotteries as exogenous shocks to investors' attention, we find: (1) individual stock returns co-move more with the market/industry returns on large jackpot days; (2) large jackpots have stronger effects on stock returns' co-movement with the market than with the industry; (3) the effect of large jackpots on return co-movement is stronger for stocks preferred by retail investors; and (4) the market under-reacts to firms' earnings surprises on large jackpot days and reverts within one week. Our findings are consistent with the existing theory that attention-constrained investors focus more on market- and sector-level information than on firm-specific information.

JEL Classification: G11, G12, G14

Keywords: limited attention, attention allocation, return co-movement, earnings surprises

I. Introduction

The concept of rational inattention dates back to the seminal work by Simon (1955, 1971) and has recently attracted increasing interest from researchers (Sims, 2003, 2010; Veldkamp, 2006; Kacperczyk, Van Nieuwerburgh, and Veldkamp 2014, 2016). Financial economists have been using the framework of attention allocation/information acquisition to explain puzzling asset pricing phenomena such as the co-movement of asset prices (Peng and Xiong, 2006; Veldkamp, 2006).[1] In particular, theoretical models predict that attention-constrained investors allocate more attention to market- and sector-level information than to firm-specific information, resulting in less firm-specific information incorporated into stock prices and higher return correlations with the market and sector indices (Peng, 2005; Peng and Xiong, 2006; Veldkamp, 2006; Veldkamp and Wolfers, 2007). Despite these clear insights and model implications, the empirical evidence on the link between investor attention and stock return co-movement is rather limited. The goal of our paper is to fill this gap in literature.

Although some proxies for attention have been proposed in the existing studies, the challenge of testing the aforementioned model implications lies in the identification of exogenous attention shocks.[2] The innovation of our paper is the use of Taiwanese nationwide lottery jackpots as repeated exogenous shocks to investors' attention, based on the premise that large jackpots distract some investors from the stock market (Gao and Lin, 2015). Consistent with previous theory, we find that stock returns co-move more with market/industry return on large jackpot days

[1] The framework is also applied to explaining home bias (Van Nieuwerburg and Veldkamp, 2009) and under-diversification (Van Nieuwerburg and Veldkamp, 2010), etc. In addition, Kacperczyk and Van Nieuwerburgh and Veldkamp (2014, 2016) test rational attention allocation from a different perspective: mutual fund managers rationally allocate attention to manage uncertainty.

[2] Proxies include Google searches (Da, Engelberg, and Gao, 2011; Liu and Peng, 2015), extreme returns (Barber and Odean, 2008), trading volume (Barber and Odean, 2008; Gervais, Kaniel, and Mingelgrin, 2001), and Hou, Peng, and Xiong, 2009), news and headlines (Barber and Odean, 2008; Fang and Peress, 2009; Fang, Peress and Zheng, 2014; Yuan, 2015), advertising expense (Chemmanur and Yan, 2009; Lou 2014), price limits (Seasholes and Wu, 2007), and online account logins (Karlsson, Loewenstein, and Seppi, 2009; Sicherman et al. 2016).

due to the distraction by lottery jackpots. In addition, large jackpots have stronger effects on stock returns' co-movement with the market than the industry. Our findings are stronger for stocks preferred by retail investors, who tend to have tighter attention constraints. Finally, we find that the market responds less to firms' earnings surprises on large jackpot days and reverts within one week. In short, attention constraints not only lead to market's negligence of firm-specific news, but they also contribute to the incorporation of more market and industry information into stock prices.

The unique features of Taiwanese nationwide lottery make it an ideal setting to demonstrate a causal link between attention shocks and return co-movement. First, the large lottery jackpots are repeated random events. The jackpot accumulates until the numbers selected by one or multiple lottery buyers match the winning numbers. Hence, the occurrence of large jackpots is unlikely to be driven by the underlying factors of stock market. Second, large jackpots have been salient nationwide events in Taiwan since the inception of the lottery in 2002 and have been shown to affect investor attention and trading behavior (Gao and Lin, 2015).[3] Third, the stock market in Taiwan is dominated by individual investors whose attention is more limited.[4]

On large jackpot days, some investors' attention on the stock market drops because they allocate more attention to these salient and exciting events in various forms. For example, they spend more time talking with friends about the lotteries. Or they are distracted by intensive media coverage on lotteries due to large jackpots. Investors thus have to be selective in processing information with the attention shocks. By focusing more on market/sector information, they could

[3] Gao and Lin (2015) argue that "*the argument is that the thrill of winning a large jackpot lottery, perhaps intensified by media coverage and investor attention around the event, substitutes some individual investors' need to trade in the stock market and thereby decreases their trading volume*." Our analysis in Section V also shows that large jackpots are highly correlated with the Google search volume index for the word "lottery" in Traditional Chinese, the official language of Taiwan.

[4] Individual investor trading volume accounts for 60% to 80% of total trading volume during our sample period.

reduce the total uncertainty of their portfolios. The incorporation of more common market information into stock prices leads to a higher correlation of stock returns with the market/industry. Hence, we propose the first hypothesis that *the co-movement of stock returns with market/industry return increases on large jackpot days*.

The extent to which market information can reduce portfolio uncertainty is different from industry information, which implies a various degree of changes in return co-movements on large jackpot days. Given Peng and Xiong (2006) predict that attention-constrained investors focus more on market information than on industry information, we propose our second hypothesis that *the change in return co-movement with market return is larger than that with industry return on large jackpot days.*

In addition, the model predictions regarding category learning in Peng and Xiong (2006) put an emphasis for retail investors. Furthermore, Gao and Lin (2015) find that retail trading volume among the stocks preferred by retail investors decreases on large jackpot days, which implies that retail investors are more affected than institutional investors by the attention shocks of large jackpots. Hence, we propose our third hypothesis that *the return on portfolios that are preferred more by retail investors correlates more closely with market return.*

Finally, when investors allocate more attention to market/industry information and less attention to firm-specific information due to distraction of large jackpots, the market may respond more slowly to new information such as firms' earnings surprises on large jackpot days than on non-large jackpot days. We thus propose our fourth hypothesis that *the market under-reacts to earnings surprises on large jackpot days, compared with earnings surprises on non-large jackpot days.*

To empirically test our hypotheses, we first define a large jackpot day dummy variable. When a lottery jackpot exceeds 500 million TWD (approximately 15 million US dollars), which is slightly higher than the 90th percentile of all jackpots throughout the sample period from January 22, 2002, to June 30, 2015, we define it as a large jackpot.[5] Next, we follow the literature and construct two measures to capture the return co-movement.[6] The first one is Pearson correlation coefficient between excess return of a single stock and the excess return of a particular category (market or sector). The second one is adjusted R-squared from estimating regressions of stock excess return on excess return of a particular category, which measures the content of less firm-specific information in the stock market. A high Pearson correlation coefficient or adjusted R-squared indicates a high level of return co-movement.

To test first hypothesis, we calculate the time-series Pearson correlation coefficients of the excess return of each individual firm and the excess return of market separately for large jackpot days and non-large jackpot days, and then obtain the difference of two coefficients for each firm. Among 817 firms, the mean and median differences in the correlation coefficients between large jackpot days and non-large jackpot days are 0.018 and 0.023, respectively. The mean of the percentage increase on large jackpot days is approximately 4% (0.018/0.445). When replacing market excess return with the excess return of the industry to which each firm belongs, the mean and median of the differences in the correlation coefficients are also significant with a smaller magnitude (0.012 and 0.017, respectively).

Similar patterns can be observed by the second measure of co-movement. We compute the adjusted R-squared by regressing each firm's excess return on the excess return of the

[5] The results are robust for alternative cutoffs, including 400 and 600 million TWD.

[6] Barberis, Shleifer and Wurgler (2005) and Peng and Xiong (2006) suggest that the R-squared value and return correlation measure the return co-movement. Morck, Yeung, and Yu (2000), Wurgler (2000), and Durnev, Morck, and Yeung (2004) argue that higher R-squared values indicate less firm-specific information in stock returns.

market/industry and then obtaining the difference for each firm. With market excess return as the independent variable, the mean and median of the differences in adjusted R-squared are 0.021 and 0.018, accounting for approximately 10% (0.021/0.211) of the mean of the adjusted R-squared on non-large jackpot days. When replacing market returns with industry returns, the mean and median of the differences in the adjusted R-squared on large jackpot days are 0.016 and 0.015, respectively, accounting for 6% (0.016/0.275) of the adjusted R-squared on non-large jackpot days. These results are consistent with our first hypothesis that stock returns co-move more with market and industry on large jackpot days when investors' attention on stocks decreases. The increase in co-movement is both statistically and economically significant.[7]

The second hypothesis examines whether investors process relatively more excess market information than excess industry information when their attention on stocks is distracted on large jackpot days. We first calculate the increases in correlation coefficients and adjusted R-squared on large jackpot days for market return and industry return, respectively. Then we obtain the differences-in-difference of Pearson correlation coefficient and adjusted R-squared for each firm. Among 817 firms, the mean and median of the differences-in-difference for the correlation coefficient are 0.005 and 0.006, respectively, and are both 0.005 for the adjusted R-squared. All statistics are significant at the 1% level and account for more than 24% of the first difference by the market category. This implies that, on average, more market information is incorporated into stock prices than industry information, lending supports for our second hypothesis.

[7] The mean of the absolute increase in return co-movement with market on large jackpot days is 0.018 (correlation coefficient) and 0.021 (adjusted R-squared). If we define the days with jackpot size bigger than 600 mil TWD as large jackpot days, the mean of the absolute change is 0.038 (correlation coefficient) and 0.043 (adjusted R-squared). The percentage change is 9% (correlation coefficient) and 20% (adjusted R-squared). For stocks traded more by retail investors, the average increase in correlation on large jackpot days can be as high as 0.1. The magnitude of the change is comparable with the effect of deletions from or additions into S&P 500 index on co-movement as shown by Barberis, Shleifer, and Wurgler (2005), and is also comparable with the effect of both being in S&P 500 index and of common mutual fund ownership on the pairwise correlation documented by Anton and Polk (2014).

To test the third hypothesis, we first construct 25 portfolios according to the ratio of individual investor's trading volume to the total trading volume over the past 22 trading days. We then use equally weighted portfolio return and market return to calculate the Pearson correlation coefficient and adjusted R-squared separately for large jackpot days and non-large jackpot days, and then obtain the difference for each portfolio. We find that as the ranking of the individual trading ratio increases, the return correlation with market on large jackpot days increases more. The sensitivity is 0.002, with a significance level of 1%, indicating that one additional increase in the ranking of retail preference leads to an increase of 0.002 in the correlation change on large jackpot days. This result is robust to alternative measures of retail investor preference, including market capitalization and idiosyncratic volatility.

For testing the fourth hypothesis, we regress the of abnormal return of earnings announcement day on the large jackpot day indicator, the decile rank of standardized unexpected earnings, the interaction of large jackpot and standardized unexpected earnings (SUE) decile rank, and control variables. The coefficient estimate on the SUE decile rank is significantly positive (0.072) and that on the interaction term is significantly negative (-0.047). The decrease in the sensitivity of abnormal returns to the SUE decile rank is approximately 65% (0.047/0.072) of the sensitivity on non-large jackpot days, which is consistent with our hypothesis that investors under-react to firm-specific news under attention shock. Moreover, within the window running from the 2^{nd} to the 5^{th} trading days after the announcement date, investors' response to earnings surprises announced on large jackpot days recovers. The effect of the large jackpot days on market response to SUE does not persist after the 5^{th} trading day.

One may concern that the large jackpot dummy do not directly capture the attention shocks. We provide corroborative evidence by using the Google search volume index for "lotto" (in

Traditional Chinese characters) as an alternative proxy to capture the attention shocks. We find that search volume is highly correlated to jackpot size and define a large search volume dummy in a similar fashion as the large jackpot dummy. The large search volume dummy variable takes the value of one if daily search volume is greater than 12.47, which is approximately 90th percentile value of the total sample. The mean and median of correlation coefficient and adjusted R-squared on large search volume days are significantly higher than those on non-large search volume days.

Additional exercises suggest that a larger attention shock, defined as a higher threshold value for categorizing a large jackpot, leads to a higher individual firms' return co-movement with the market/industry. Specifically, we repeat the analytical procedures with two alternative lottery jackpot cutoffs: 400 and 600 million TWD. The mean and median changes of correlation coefficients and the adjusted R-squared increase significantly with the cutoff values. This result indicates that more categorical information is incorporated into stock prices when investors face a tighter attention constraint.

Our paper contributes to the literature by providing a clean identification on testing the relationship between attention allocation and return co-movements. It thus differs from and complements Liu and Peng (2015) who analyze how attention, measured by Google Search volume, is affected by macro news in two aspects. First, we investigate the effect of attention shock on the composition of firm-specific and category-level information into stock prices and return co-movement while their focus is on how information shock affects the attention allocation. Second, the exogenous shocks of large jackpot lotteries on investor attention help us to draw a causal inference. That being said, both Liu and Peng (2015) and our work provide complementary evidence to the model implications of Peng and Xiong (2006).

Our paper is also related to research on the relationship between return synchronicity and price informativeness. We add to this strand of literature by demonstrating that the return synchronicity increases with a decline in firm-specific information. Moreover, we also illuminate the literature regarding the determinants of stock return co-movement.[8] Our results provide a new perspective to explain return co-movement: the category learning from market/industry information caused by rational inattention.

The rest of the paper is organized as follows. The next section describes Taiwanese lottery data and the construction of major variables. Section III presents our main results on return co-movement with market information, with industry information, the comparison between learning from market information and learning from industry information, and how retail investor ownership affects co-movement. Section IV addresses the argument that less firm-specific information is incorporated into stock prices due to attention constraints by examining stock market reactions to earnings announcements on large jackpot days. Section V conducts robustness checks, and Section VI concludes.

II. Data and Descriptive Statistics

A. The Taiwanese Lottery and Large Jackpots

Our analysis combines three nationwide lotteries in Taiwan: Lotto, Big Lotto, and Super Lotto (Powerball). Lotto began on January 22, 2002, and was replaced by Super Lotto on January 24, 2008. Big Lotto began on January 5, 2004, and continues to the present. Super Lotto began on January 24, 2008, and continues to the present. The website of the Taiwan Lottery Company

[8] Sources of co-movement include cash flows or discount rates (traditional theory), category view (Barberis and Shleifer, 2003), habitat view, information diffusion view, friction- or sentiment-based view (Barberis, Shleifer and Wurgler, 2005; Kumar, Page and Spalt, 2013), gambling (Kumar, Page and Spalt, 2016), and etc.

reports the cumulative prize of each lottery drawing as its jackpot size.[9] The jackpot continues to increase until the prize is won. Table 1 presents descriptive statistics of the lottery jackpot size. The sample period is from January 22, 2002 to June 30, 2015, which includes a total of 2,654 lottery drawings. We define a large jackpot to be a cumulative lottery prize (Lotto, Big Lotto, or Super Lotto) of greater than 500 million TWD (approximately 15 million US dollars), which is slightly greater than the 90th percentile value of the total sample. In total, there are 341 large jackpot lotteries greater than 500 million TWD.

[Table 1 about here]

As stated by Gao and Lin (2015), a large lottery jackpot "draws their [investors'] attention away from trading stocks" and leads to "investors reallocating their attention toward activities that are fun and exciting". Investors allocate less attention to the costly information in the stock market and instead allocate more attention to processing lottery information. To verify this intuitive argument, we compare the daily Google search volume index with lottery jackpot size as Google search volume can be viewed as a direct measure of attention (Da, Engelberg, and Gao, 2011). The sample period is from January 4, 2004, to June 30, 2015. We manually collect the Google weekly search index over the sample period and the Google unadjusted daily search index over each quarter. The search word is "樂透" which means "lottery" or "lotto" in English, and the search region is Taiwan.[10] Since unadjusted daily search index is the relative search volume over each

[9] The data for lotteries including Lotto, Big Lotto, and Super Lotto were collected manually from the Taiwan Lottery Company (http://www.taiwanlottery.com.tw/eng/about/tlc.asp).
[10] Google search volume data were collected manually from Google Trends (https://www.google.com/trends/).

quarter, we calculate adjusted daily search index as (weekly search index to which the day belongs) * (ratio of unadjusted daily search index to the weekly average of unadjusted daily search index).

Appendix Figure A1 is provided to show that the Google daily search volume index experiences an increase on each large jackpot day. Specifically, Google search volume index is significantly higher on large jackpot days than on non-large jackpot days. The mean difference is 11.961, which is approximately 1.5 times the mean of daily search volume on non-large jackpot days (Panel A of Appendix Table A1). A large Google search volume indicator is defined as a daily search volume index greater than 12.47, which is approximately 90th percentile value of the total sample. Panel B in Appendix Table A1 suggests that daily Google search volume is positively correlated with lottery jackpot size. The indicator of large jackpot and indicator of large search volume are also highly correlated. This result supports the assumption that investors' attention is attracted by lottery large jackpot, thus leaving less attention in the stock market.[11]

Because large jackpots occur only due to the absence of winners in several previous lottery drawings, large jackpots are not likely to be driven by macro-economic factors that affect the stock market. Occurrences of large jackpots can be regarded as repeated natural experiments that capture shocks to the attention of investors.

B. Stock Returns, Individual Investor Trading Ratio, and Earnings Announcements

Stock trading data are from the Taiwan Economic Journal. We focus on stocks listed on the Taiwan Stock Exchange (TWSE). Firms with less than 100 trading days in the sample period (4 firms) and firms with first listing day later than the year 2012 (48 firms) are excluded. The final sample includes 817 stocks listed on the TWSE. The TWSE provides daily returns for individual firms, industry indices, and the market index, which can be used to test individual firm return co-

[11] Results from using Google weekly search volume are similar to results from using Google daily search volume.

movement with market return or industry return. Before July 2007, 20 industry sectors were classified by the Taiwan Stock Exchange Corporation.[12] After 2007, the chemical industry and the electronics industry were sub-classified into additional types, resulting in 29 industry sectors.[13] The corresponding index returns are available from that time onward. To obtain the daily excess returns, we calculate the daily risk-free rate from the rate on central bank one-month time deposits in Taiwan.

The TWSE provides the daily number of shares bought and sold by institutional investors. We follow the formula constructed by Gao and Lin (2015) to estimate the daily buy and sell volume of each stock by retail investors:

$$V_{i,t} = TV_{i,t} - (IB_{i,t} + IS_{i,t}) \tag{1}$$

$TV_{i,t}$ is the total shares of firm i traded on day t; $IB_{i,t}$ is the number of shares of firm i bought by institutional investors on day t; and $IS_{i,t}$ is the number of shares of firm i sold by institutional investors on day t. The assumption is that the trading counterparties of an institutional buy or sell are individuals.

In the analysis of the market reaction to earnings announcements, we define an earnings surprise by the SUE according to Chan, Jegadeesh, and Lakonishok (1996) and Chordia and Shivakumar (2006). The sample period is from 2002 to 2015. After the SUE and all control variables are calculated, there are 948 earnings announcement days in our sample, 81 of which are large jackpot days. In total, there are 21,688 firm-date data points, and 1,854 of them announce earnings on large jackpot days.

[12] The industry sectors are cement, food, plastic, textile, machinery, wire & cable, chemical, glass & ceramic, paper, steel, rubber, auto, electronics, construction, transportation, tourism, financials, retail, gas & oil, and others.

[13] Firms in the chemical industry are sub-divided into chemical and biotech firms; firms in the electronics industry are sub-divided into semiconductor, computer, optoelectronics, communications, electronic components, electronic channel, information, and other firms.

III. Main Results for Return Co-Movement

This section empirically tests the hypotheses that (1) on average, individual firms' return co-movement with the market and their corresponding industry increases when investors' attention to the stock market decreases due to large lottery jackpots; (2) the increase in individual firms' return co-movement with the market is stronger than that with the industry on large jackpot days; and (3) the effect of large jackpots on the change in return co-movement with the market is stronger for stocks preferred by retail investors. The hypotheses originate from the model developed by Peng and Xiong (2006). Given the existence of limited attention and the vast amount of information, investors must be selective in processing information. Attention-constrained investors allocate more attention to market- and sector-level information and less attention to firm-specific information. Under realistic conditions, "investors allocate more attention to the market factor than to a sector factor and more attention to a sector factor than to a firm-specific factor" and "the return correlation between any two firms is higher".[14]

A. Return co-movement with the market

The first measure of return co-movement is the time-series Pearson correlation coefficient. To capture the change in the correlation coefficient due to the occurrence of large jackpots, we separate the return data into a large jackpot group and a non-large jackpot group. The large jackpot group contains firm-day stock returns when the jackpot size is larger than 500 million TWD on the day. For any firm *i*, the correlation between the firm excess return and the market excess return on non-large jackpot days is as follows:

$$Corr_i^{NLJD} = \frac{\sum_t (Ret_{i,t} - \overline{Ret}_i)(MktRet_t - \overline{MktRet})}{\sqrt{\sum_t (Ret_{i,t} - \overline{Ret}_i)^2} \sqrt{\sum_t (MktRet_t - \overline{MktRet})^2}} \quad (2)$$

[14] Peng and Xiong (2006)'s Section 4, Proposition 1 and 3.

where *t* are non-large jackpot days. Similarly, the correlation between firm excess return and market excess return on large jackpot days is as follows:

$$Corr_i^{LJD} = \frac{\sum_T (Ret_{i,T} - \overline{Ret}_i)(MktRet_T - \overline{MktRet})}{\sqrt{\sum_T (Ret_{i,T} - \overline{Ret}_i)^2}\sqrt{\sum_T (MktRet_T - \overline{MktRet})^2}} \quad (3)$$

where *T* are large jackpot days. For analyses throughout this paper, $Ret_{i,t}$ is the excess return of firm *i* at time *t*; $MktRet_t$ is the market excess return at time *t*. $\overline{Ret}$ and $\overline{MktRet}$ are the sample mean. Then, we obtain $Corr_i^{LJD}$-$Corr_i^{NLJD}$ for any firm *i*. Reported in Table 2 are the mean and median of $Corr_i^{LJD}$, $Corr_i^{NLJD}$ and $Corr_i^{LJD}$-$Corr_i^{NLJD}$ among all 817 firms.

The second measure of the return co-movement is the adjusted R-squared obtained from the market model. After separating the return data into a large jackpot group and a non-large jackpot group, for firm *i*, we perform the following regression on non-large jackpot days *t*:

$$Ret_{i,t} = \alpha_i + \beta_i * MktRet_t + \varepsilon_{i,t} \quad (4)$$

Thus, we obtain the adjusted R-squared for firm *i* on non-large jackpot days $\bar{R}_i^{2\ NLJD}$. The same regression for large jackpot days *T* yields the adjusted R-squared for firm *i* on large jackpots days $\bar{R}_i^{2\ LJD}$. Table 2 displays the mean and median of $\bar{R}_i^{2\ LJD}$, $\bar{R}_i^{2\ NLJD}$, and $\bar{R}_i^{2\ LJD} - \bar{R}_i^{2\ NLJD}$ among all 817 firms.

[Table 2 about here]

On average, the correlation between an individual firm's return and market return increases on large jackpot days. The mean and median of the correlation coefficients calculated from large jackpot days exceed those calculated from non-large jackpot days. The mean and median differences are 0.018 and 0.023, respectively, with a significance level at 1%. In terms of the adjusted R-squared, the mean and median changes are 0.021 and 0.018, respectively, and significant. To test the significance of the mean and median of the differences, we use both paired t test and Wilcoxon signed-rank test which is a nonparametric test without assuming the population distribution.

B. Return co-movement with the industry

Investors also learn sector-wide information in the stock market when they rationally allocate a certain share of total attention to processing lottery information due to the exogenous occurrences of large jackpots. The Taiwan Stock Exchange Corporation provides the industry index return within our sample period. The annual frequencies of firms in these sectors are reported in Appendix Table A2. We adopt the same return co-movement measures as in Subsection III.A, except that $MktRet_t$is replaced by $IndRet_{i,t}$. $IndRet_{i,t}$is the index excess return of the industry to which firm *i* belongs. The mean and median of $Corr_i^{LJD}$, $Corr_i^{NLJD}$, $Corr_i^{LJD}$-$Corr_i^{NLJD}$, $\bar{R}_i^{2\ LJD}$, $\bar{R}_i^{2\ NLJD}$, and $\bar{R}_i^{2\ LJD} - \bar{R}_i^{2\ NLJD}$are presented in Table 3.

[Table 3 about here]

The mean and median changes in the correlation coefficients on large jackpots are 0.012 and 0.017, respectively; the mean and median changes in the adjusted R-squared are 0.016 and 0.015, respectively. All differences are significant at a level of 1%. Because the industries were sub-divided in 2007, the results are similar to those in Table 3 if we use the old (broader) industry index return throughout the sample period. We also analyze the correlation and the adjusted R-squared increase on large jackpots separately for all industry sectors. Appendix Table A3 shows the results. The mean and median changes in the return co-movement are significant and positive in 10 industries in which the number of firms represents approximately 50% of all firms. The results in the previous section and this section are supportive to our first hypothesis that stock return co-moves more with the market and industry indices when investors' attention is distracted.

C. Comparison between co-movement with the market and with the industry

The second hypothesis is that the excess co-movement of individual firms' return with market return on large jackpot days is stronger than that with industry return. To test, we compare the differences-in-difference in the correlation coefficient (adjusted R-squared) on large/non-large jackpot days and using market/industry as the category. Specifically, model (2) follows the procedures in Subsection III.A, that is, individual firms' return co-movement with the market and for any firm *i* among the 817 firms. We have the following:

$$\begin{cases} Corr(2)_i^{LJD}, Corr(2)_i^{NLJD}, Corr(2)_i^{LJD} - Corr(2)_i^{NLJD} \\ \overline{R(2)}_i^{2\ LJD}, \overline{R(2)}_i^{2\ NLJD}, \overline{R(2)}_i^{2\ LJD} - \overline{R(2)}_i^{2\ NLJD} \end{cases} \tag{5}$$

Model (1) conducts the same analyses in Subsection III.B, that is, individual firms' return co-movement with the industry and for any firm *i* among the 817 firms. We obtain the following:

$$\begin{cases} Corr(1)_i^{LJD}, Corr(1)_i^{NLJD}, Corr(1)_i^{LJD} - Corr(1)_i^{NLJD} \\ \overline{R(1)}_i^{2\ LJD}, \overline{R(1)}_i^{2\ NLJD}, \overline{R(1)}_i^{2\ LJD} - \overline{R(1)}_i^{2\ NLJD} \end{cases} \tag{6}$$

To compare the co-movement changes between the two models, we calculate the difference between models for firm *i* as follows:

$$\begin{cases} [Corr(2)_i^{LJD} - Corr(2)_i^{NLJD}] - [Corr(1)_i^{LJD} - Corr(1)_i^{NLJD}] \\ [\overline{R(2)}_i^{2\ LJD} - \overline{R(2)}_i^{2\ NLJD}] - [\overline{R(1)}_i^{2\ LJD} - \overline{R(1)}_i^{2\ NLJD}] \end{cases} \tag{7}$$

Table 4 shows the mean and median of the above statistics among 817 firms. The last column shows that the mean and median of the model differences (the correlation with the market minus the correlation with the industry) are 0.005 and 0.006, respectively, and the mean and median of the model differences (adjusted R-squared) are both 0.005. All estimates are significant at the 1% level.[15] The result is consistent with our second hypothesis.

[Table 4 about here]

Summarizing the findings, on days when the lottery jackpot exceeds 500 million TWD, because of the thrill of a chance to win a large amount of money or the higher subjective expected return from lotteries, investors replace their attention to the stock market with attention to the lottery. Attention-constrained investors will make pricing and investment decisions by processing more market- and sector-wide information. The category-level information, including market and sectors, is incorporated in the stock returns. Therefore, individual firm' return co-movement with the market and the sector will increase. In addition to using mean/median of changes in correlation and R-squared as a measure of return co-movement, we test the pair-wise correlations among firms in the same market or industry. Because more common factors are priced in the stock returns on large jackpot days, the pair-wise correlation among firms is also higher (Appendix Table A5).[16] Furthermore, because market-level information is relevant for more stocks and can be used to lower the uncertainty of portfolios, the return co-movement with the market is stronger than that with the industry.

D. Individual investor ownership and stock returns' co-movement

[15] The comparison between investors' learning from market information and from industry information within industries is provided in Appendix Table A4. In approximately 10 industries, the co-movement with the market is larger than the co-movement with the industry.

[16] This finding supports the Proposition 3 of Peng and Xiong (2006) and mitigates the concern that increase in co-movement on large jackpot days is driven by the leaving of noise gamblers who tend to have correlated trading (Barber, Odean, and Zhu 2009a; Barber, Odean, and Zhu 2009b).

This subsection tests the third hypothesis that the returns of stocks preferred by individual investors co-move more with the market/sector than those preferred by institutional investors. The attention of retail investors is more likely to deviate from the stock market to the lottery market because they are more likely to treat stocks as lotteries (Gao and Lin, 2015).

From daily observations, we sort all stocks into 25 portfolios according to the average individual trading ratio in the past 22 trading days.[17] The individual trading volume is calculated as in equation (1). The individual trading ratio is defined as the past 22 days' total individual trading volume divided by the past 22 days' total trading volume:

$$Itr_{i,t} = \frac{\sum_{t-22}^{t-1} V_{i,t}}{\sum_{t-22}^{t-1} TV_{i,t}} \tag{8}$$

We construct 25 portfolios, denoted as portfolio *j*, where *j*=1, 2, …, 25 is the ranking of the individual trading ratios (smallest individual trading ratio: *j*=1). The time-series equally weighted portfolio returns are separated into a large jackpot group and a non-large jackpot group. The time-series Pearson correlation between the portfolio return and the market return for portfolio *j* on non-large jackpot days *t* and large jackpot days *T* are as follows:

$$Corr_j^{NLJD} = \frac{\sum_t (PortRet_{j,t} - \overline{PortRet}_j)(MktRet_t - \overline{MktRet})}{\sqrt{\sum_t (PortRet_{j,t} - \overline{PortRet}_j)^2}\sqrt{\sum_t (MktRet_t - \overline{MktRet})^2}} \tag{9}$$

$$Corr_j^{LJD} = \frac{\sum_T (PortRet_{j,T} - \overline{PortRet}_j)(MktRet_T - \overline{MktRet})}{\sqrt{\sum_T (PortRet_{j,T} - \overline{PortRet}_j)^2}\sqrt{\sum_T (MktRet_T - \overline{MktRet})^2}} \tag{10}$$

$PortRet_{j,t}$ denotes the equally weighted excess return for portfolio *j* at time *t*. The second measure of portfolio return co-movement is the adjusted R-squared from the following regressions. For portfolio *j*, the time-series regression for non-large jackpot days *t* and large jackpot days *T* is as follows:

[17] We also group stocks by the current day individual trading ratio and the results are very similar.

$$PortRet_{j,t} = \alpha_j + \beta_j * MktRet_t + \varepsilon_{j,t} \quad (11)$$

$$PortRet_{j,T} = \alpha_j + \beta_j * MktRet_T + \varepsilon_{j,T} \quad (12)$$

For any portfolio j, $Corr_j^{LJD} - Corr_j^{NLJD}$ and $\bar{R}_j^{2\ LJD} - \bar{R}_j^{2\ NLJD}$ are calculated. In an attempt to compare the market learning among portfolios, we perform the following OLS regressions, where *j* is the ranking of portfolios according to individual trading ratio:

$$Corr_j^{LJD} - Corr_j^{NLJD} = \alpha_j + \beta * \mathrm{J} \quad (13)$$

$$\bar{R}_j^{2\ LJD} - \bar{R}_j^{2\ NLJD} = \alpha_j + \beta * \mathrm{J} \quad (14)$$

The regression results are displayed in Table 5. The coefficient on ranking *j* is positive and significant at 1% level, which means that the return of portfolios with a higher individual trading ratio co-moves more with the market. Specifically, one additional increase in the ranking of individual trading ratios drives up the correlation coefficient and adjusted R-squared by 0.002.

[Table 5 about here]

To provide a clearer presentation of the results, we sort stocks into 10 portfolios, and the analyses are repeated as above. Figure 1 shows how individual investor ownership affects the return correlation with the market and adjusted R-squared from the market model. The ranking of individual trading ratios is on the x-axis, and the increase in the correlation coefficient (Panel A) or adjusted R-squared (Panel B) on large jackpot days is on the y-axis. With the increase in ranking, the return co-movement generally increases.

[Figure 1 about here]

In addition to the direct measure of retail investor preference, which is the individual trading ratio, by using two alternative measures of retail investor preference, we get the similar results.[18] Therefore, greater change in return co-movement with the market for stocks preferred by retail investors on large jackpot days is robust to different measures of retail investors' stock-picking preferences. Because retail investors are more likely to reduce their total attention on stock market information when attracted by large lottery jackpots, the category learning reflected in the return co-movement with the market/sector is stronger for stocks preferred and traded by these investors.

IV. Market Reaction to Earnings Announcements

To further support our claim that fewer firm-specific factors are contained in stock returns, we test our forth hypothesis that market under-reacts to earnings surprises on large jackpot days. The under-reaction to earnings surprises indicates that investors pay less to inattention to firm-specific news on large jackpot days.

Hirshleifer, Lim, and Teoh (2009) show that the number of same-day earnings announcements by other firms is correlated to an immediate price and volume under-reaction to the firm's earning surprise and a stronger post-announcement drift. They attribute this phenomenon to the channel of increased information load faced by investors. Similarly, DellaVigna and Pollet (2009) find that investor inattention affects stock prices by comparing the market responses to earnings surprises released on Fridays and that on other days. Their results indicate a less immediate response to and more drift for earnings surprises released on Fridays.

[18] The two additional measures of retail investors' preferences are (1) market capitalization and (2) idiosyncratic return volatility. Retail investors prefer small stocks (Kumar and Lee, 2006), and stocks with high idiosyncratic return volatility (Barberis and Huang, 2008; Kumar, 2009; Green and Hwang, 2012). Using the same methodology as in Table 5, we present the regression results in Appendix Table A6.

The same-day earnings announcements or Fridays' announcements, however, may suffer from the endogeneity concern. Particularly, firms with certain traits (for instance, poor operations) may choose to announce their earnings on Fridays (or, in certain cases, on days when many other firms make the announcements) to mitigate stock price fluctuations or to meet some deadline. The under-reaction and drift may be driven by those characteristics rather than investor or analyst inattention. One unique advantage of using large jackpots in our paper is that large jackpot days occur randomly and earnings announcement dates are pre-determined, so it is unlikely that any firm can exploit the timing of earnings announcements. We can safely attribute the market under-reactions toward earnings surprises to investor inattention on firm-specific information.

A. Under-reaction of stock returns to earnings announcements

We use the quarterly earnings per share (EPS) data announced by individual firms to test our hypothesis. The earnings surprise is measured by the standardized unexpected earnings (SUE). Following the methodology used by Chan, Jegadeesh, and Lakonishok (1996) and Chordia and Shivakumar (2006), the SUE of firm *i* at quarter *q* is calculated as follows:

$$SUE_{i,q} = \frac{Eps_{i,q} - Eps_{i,q-4}}{\sigma_{i,q}} \qquad (15)$$

Eps_q is the earnings per share at quarter q, and Eps_{q-4} is the earnings per share at the same quarter in the previous year. $\sigma_{i,q}$ is the standard deviation of $Eps_{i,q} - Eps_{i,q-4}$ in the previous eight consecutive quarters. The SUEs are winsorized at the 1st and 99th percentiles to minimize the influences of outliers. The breakpoints for the SUE deciles are decided for each fiscal quarter. For the smallest 10% of SUE values, the SUE decile rank is equal to 1; for the largest 10% of SUE values, the SUE decile rank is equal to 10. The SUE decile rank is used to mitigate the forecast error because the relationship between stock returns and the SUE is likely to be nonlinear (Kothari, 2001). We estimated betas from the market model from 300 trading days to 46 trading days before the date of the announcement. The announcement day abnormal return is as follows:

$$AR_{i,t} = Ret_{i,t} - \overline{\beta}_{i,t} MktRet_t \tag{16}$$

The cumulative abnormal return for firm *i* is the buy-and-hold abnormal return.

$$CAR(\tau + h, \tau + H)_{i,t} = [\prod_{j=\tau+h}^{\tau+H} (1 + Ret_{i,j})] - 1 - \overline{\beta}_{i,t} [\prod_{j=\tau+h}^{\tau+H} \left(1 + MktRet_j\right) - 1] \tag{17}$$

To capture the effects of a large jackpot on the market reactions to earnings surprises, we perform regressions of announcement day abnormal return and cumulative abnormal returns within the windows of (0,1), (2,5) and (6,60) on the large jackpot dummy, the SUE decile rank, the interaction term large jackpot dummy*SUE decile rank, and a set of control variables:

$$CAR_{i,t} = \alpha_0 + \alpha_1 LJD_t + \alpha_2 SUE\ decile\ rank_{i,t} + \alpha_3 \left(LJD_t * SUE\ decile\ rank_{i,t}\right) + \sum_{k=1}^{K} c_k Controls_{i,t} + \varepsilon_{i,t} \tag{18}$$

The controls include industry dummies, the decile of market capitalization, the number of announcements at announcement day, the past 22 days' average individual trading ratio, the past 22 days' average return, the past 22 days' idiosyncratic volatility before the announcement day, the year, the month, the day of the week, and turnover. Standard errors are adjusted for heteroskedasticity and clustered by both the day of announcement and firm. Hypothesis 4 posits that announcement-day stock returns are less sensitive to earnings surprises on large jackpot days. Thus we expect α_3 to be negative.

Panel A of Table 6 shows that the estimates of α_2 are positive and significant at the 1% level, which means that, on average, one additional increase in the SUE decile leads to an abnormal return and CAR (0,1) increase of 0.08 and 0.211 without controls (0.072 and 0.217 with controls), respectively, for non-large jackpot days. The estimates of α_3 are negative and significant (for announcement-day abnormal return and CAR(0,1): -0.056 and -0.096 without controls; -0.047 and -0.089 with controls). All returns are in percentage form. The implication is that market reactions to earnings surprises are significantly less sensitive on large jackpot days by 0.047/0.072=65% (column 2), compared with that of non-large jackpot days. The result is in line with our forth hypothesis and supports our main claim that there is less firm-specific information incorporated into stock prices on large jackpot days.

[Table 6 about here]

B. Response of stock returns to earnings surprises after earnings announcements

Because the market under-reacts to earnings surprises on large jackpot days, we expect market responds to earnings surprises in the following few days when investor attention to stock-specific information resumes. When using CAR (2,5) as the dependent variable, in addition to the standard controls, we also include the number of large jackpots during the window of 2 to 5 divided by 4 and interaction of that number with the SUE decile rank to control for the possibility that consecutive large jackpot days slow the resuming process. We find that, from the second trading day after the announcements, the sensitivity of cumulative stock returns to the SUE decile rank is stronger for earnings announced on large jackpot days. The regression results with CAR (2,5) as the dependent variable are shown in Panel B of Table 6. With controls, one additional increase in the SUE decile rank will result in an increase of 0.197 for CAR (2,5); for stocks whose earnings announcements are made on large jackpot days, there is an additional increase of 0.137 to the cumulative abnormal return. This finding suggests that lagged market reactions to earnings surprises are significantly stronger on large jackpot days by 0.137/0.197=70% (column 2) compared with non-large jackpot days.

A natural question is what happens after the 5th trading day (approximately one week later)? Does the impact of a large lottery jackpot on the announcement day on the stock market reaction to earnings surprises still persist? The results in columns (3) and (4) help to address these questions. We perform the same regression as in the previous subsection, except for that the dependent variable is CAR (6,60). Additional control variables are the number of large jackpot days within the window divided by the number of trading days within the window and its interaction with the SUE decile rank. With one increase in the SUE decile rank, the cumulative abnormal return from the 6th trading day to the 60th trading day increases by 0.458 without controls and 0.496 with controls. However, the coefficient on the interaction term of the large jackpot dummy and the SUE decile rank is no longer significant. The interpretation is that the effect of large jackpots on market response to earnings surprises evaporates after one week. Because of the delayed response to earnings surprises on large jackpot days over the window of (2,5), the influence slightly and negatively reverts. Ultimately, a large jackpot on the announcement day does not affect market reactions to earnings surprises permanently.

V. Robustness Checks

A. Capturing investor attention shocks using Google search volume on "lottery"

In addition to lottery jackpot size, the alternative proxy we use to capture shocks to investor attention is the search volume of "lottery" on Google, which has been verified to be highly correlated with lottery jackpots in Appendix Table A1.[19] As a robustness check of the effect of attention shocks to investors on return co-movement with market/industry, we conduct similar analyses with Tables 2, 3 and 4, by replacing large/non-large jackpot days with large/non-large search volume days. Large search volume days are defined as those on which the daily search volume index of "lotto" (in Traditional Chinese) on Google is greater than 12.47, which is approximately 90th percentile value of the total sample from January 4, 2004, to June 30, 2015.

[Table 7 about here]

Following the same methodology in Table 2, we find that the mean and median of correlation coefficient/adjusted R-squared between individual excess return and market excess return on large search volume days are significantly higher than those on non-large search volume days. The mean and median differences are 0.026 and 0.03 respectively for correlation coefficients, and 0.023 and 0.025 for adjusted R-squared. Similarly, mean and median of correlation coefficient/adjusted R-squared between excess stock return and corresponding industry excess return on large search volume days are significantly higher than those on non-large search volume days. The mean and median differences are each 0.02 for correlation coefficients, and 0.02 and 0.018 for adjusted R-squared.

We also compare learning from market information with learning from industry information using the method discussed in Table 4. Results in last column of Table 7 show that more market information is incorporated into stock prices than industry information when there is

[19] Schmidt (2013) uses Google search for different sports as a proxy for attention shocks and find that 4 countries show significantly negative changes in synchronicity and 13 countries have positive changes in synchronicity. The mixed evidence is also subject to endogeneity concerns such as seasonality which is shown by Liu and Peng (2015).

bigger search interest of "lottery" on Google. The mean and median differences-in-difference are 0.005 and 0.008, respectively, for correlation coefficients, and 0.003 and 0.005, respectively, for adjusted R-squared. This alternative measure of attention shocks verifies our first two hypotheses that more categorical information is incorporated into stock prices when investors have tighter attention constraints, particularly for market information. Moreover, using the same methodology with Tables 5, 6, and 7, we find that portfolios preferred by retail investors co-move more with market on large Google search volume days, and that investors under-react to earnings surprises on large Google search volume days, which supports the third and fourth hypotheses.

B. Gambling effect

One may concern that large jackpots could be correlated with gambling sentiment, which would drive up the return co-movement (Kumar, Page and Spalt, 2016). We need to notice that potential increase of gambling sentiment on large jackpot days could not occur without the attention shocks. Although these two are difficult to disentangle, we provide attempts to confirm that the increase in co-movement on large jackpot days is not driven merely by gambling effect. If the co-movement change is driven by gambling sentiment, the effect should be concentrated on lottery-type stocks. But we find that the increase in return co-movement with the lottery-type portfolio return is similar to the increase in return co-movement with non-lottery-type stocks (Appendix Table A7); and that stocks more preferred by retail investors co-move more with market, excluding the effect of lottery-type features.

In order to tease out lottery-type features, we calculate average closing price in the past 22 trading days, idiosyncratic skewness and idiosyncratic volatility in the past 150 trading days for each firm-date observation, which are considered as features preferred by gamblers (Kumar, 2009). Each date, we perform cross-sectional regressions of individual trading ratio (past 22 trading days) on average closing price (past 22 trading days), idiosyncratic skewness and idiosyncratic volatility (past 150 trading days), and then obtaining residuals, which represent characteristics preferred by retail investors other than lottery-type features. Then we sort stocks into 25 portfolios by the residual (ranking of retail investor preference=1 if the residual falls in the lowest 4% group) and calculate the equally weighted portfolio excess returns. After calculating the increase in co-movement with market for each portfolio on large jackpot days similar to Table 5, we regress the increase in co-movement on the ranking of retail preference. Table 8 shows that there is a significant increase in the excess co-movement with market if the portfolio is more preferred by retail investors net of the lottery-type features.

[Table 8 about here]

Following the same methodology in Figure 1, for better presentation of the results, we present Figure 2. We sort stocks into 10 portfolios by the residual (ranking of retail investor preference=1 if the residual falls in the lowest 10% group) and calculate the equally weighted portfolio excess returns. Then the increase in co-movement with market on large jackpot days is calculated for each portfolio. The increase in Pearson correlation/adjusted R-squared between excess return of each portfolio and market excess return on large jackpot days is increasing with retail investor trading ratio, net of the lottery-type features. [20] Excluding the effect of lottery-type features, the return of portfolio preferred by retail investors who experience larger attention constraints on large jackpot days co-move more with the market return. Thus the effect of the

[20] As in Table 8, we also run the regression of increase in each portfolio's return co-movement with market on the ranking of retail preference for each portfolio. With the increase in ranking of retail preference excluding lottery-type features, there is a significant increase in the excess return co-movement.

increase in co-movement on large jackpot days is not likely to concentrate on lottery-type stocks only.

[Figure 2 about here]

C. Alternative cutoffs for defining large jackpots

For the previous analyses, we adopt a cumulative lottery prize (jackpot) of 500 million TWD as the cutoff to determine whether a day is a large jackpot day because it is approximately the 90th percentile of all lottery jackpots. When the jackpot is above this value, we assume that the attention of investors, particularly among retail investors, will allocate more attention on market- and sector-level information and less firm-specific information. As one set of robustness checks, we provide individual firms' return co-movement with the market on large jackpot days compared with non-large jackpot days by alternative reasonable cutoffs, 400 and 600 million TWD. For cutoffs of 400, 500, and 600 million TWD, there are 429, 304 and 215 large jackpot trading days, respectively.[21] It is expected that, with the increase in the large jackpot cutoffs, the attractiveness of the lottery is increased on large jackpot days, and less attention is devoted to the stock market. Investors rationally learn more information from the category and less from firm-specific news, which is reflected in the higher co-movement with the market/industry due to the larger jackpot cutoffs.

Following the same methodology in Subsection III. A, Table 9 reports the mean and median of the differences in the Pearson correlation coefficient and adjusted R-squared on large

[21] We do not adopt 700 million as a cutoff because the number of large jackpot trading days with this threshold is only approximately half of that with the 500 million threshold and is too small for our analysis.

and non-large jackpot days. When the cutoff value is 400 million TWD, the means of changes in the correlation and adjusted R-squared are 0.008 and 0.011, respectively; when the cutoff value is 500 million TWD, the means of the changes in the correlation and adjusted R-squared are 0.018 and 0.021, respectively; and when the cutoff value is 600 million TWD, the means of the changes in the correlation and adjusted R-squared are 0.038 and 0.043, respectively. It is clear that the magnitude of co-moving with the market increases with larger cutoff values. There are similar trends for the median of changes in the correlation and adjusted R-squared.

[Table 9 about here]

In addition, the return co-movement with the industry has been analyzed in detail and can be presented upon request. To better understand the co-movements based on different lottery jackpot cutoffs, we present Figures 3 and 4. Panel A focuses on the co-movement measure of the Pearson correlation coefficients; Panel B emphasizes the second measure, the adjusted R-squared. The figure on the left records the means of the changes in the correlation; the figure on the right shows the medians of the changes in the correlation. The x-axis reports the value of the three alternative cutoff values for a large jackpot; the y-axis marks the difference in co-movement between large jackpot days and non-large jackpot days.

[Figure 3 and Figure 4 about here]

The conclusion is that the means and medians of co-movement with both the market and the industry increase with larger lottery jackpots. Larger lottery jackpots lead to larger increases in attention allocated to the lottery and decreases in attention allocated to stocks. Reflecting return co-movement, category learning with the market and with the industry strengthens. The lines in Figure 4 show the comparison between learning from the market (the broadest category) and from the industry (a less broad category). There is a significant increase in the difference between learning from the market and learning from the industry with larger jackpot cutoff values.

VI. Conclusion

Attention is a scarce cognitive resource (Kahneman, 1973). The previous theory predicts that investors tend to allocate more attention to market/industry level information relative to firm-specific information due to attention constraints. In this paper, we test how investors reallocate their limited attention and how this reallocation is reflected in the co-movement of stock prices with a particular category. By using random occurrences of large lottery jackpots in Taiwan to capture repeated and nationwide exogenous shocks to investor attention capacity, we provide empirical evidence to the predictions proposed by theoretical papers that attention-constrained investors process more market- and sector-wide information and less firm-specific information. Specifically, more market- and sector-wide information and less firm-specific information are incorporated into stock prices during the large jackpot days than non-large jackpot days.

We define a large jackpot as one that exceeds 500 million TWD. Compared with non-large jackpot days, we find that on large jackpot days: (1) stock return co-movement with the market/industry return increases; the average increase in Pearson correlation coefficient (adjusted R-squared) with the market is 4% (10%), and the average increase in the correlation (adjusted R-squared) with the industry is 2.5% (6%); (2) the increase in return co-movement with the market return is larger than that with industry return. The difference is at least 24% of the increase in the return co-movement with the market; (3) the increase in co-movement is more prominent for stocks preferred by retail investors; and (4) announcement-day stock abnormal returns under-react to earnings surprises announced on large jackpot days by 65%, and the effect of a large jackpot on the return sensitivity to the SUE decile rank decreases (70%) between the 2nd and 5th trading days.

By providing a clean identification of exogenous and repeated attention shocks, our paper contributes to the literature by confirming the predictions of theoretical models that attention-constrained investors allocate more attention to market- and sector-level information than to firm-specific information, resulting in higher return correlations with the market/sector. It also adds to literature about the relationship between return synchronicity and price informativeness by demonstrating that the return synchronicity increases with a decline in firm-specific information. Moreover, our results provide a new perspective to explain return co-movement, namely, the category learning from market/industry information caused by rational inattention. Although the settings of the Taiwanese evidence cannot be completely replicated for other markets, we expect that the rational allocation of attention to category-level information is universal.

References

Anton, M., Polk C., 2014. Connected stocks. Journal of Finance 69, 1099–1127.

Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and News on the buying behavior of individual and institutional investors. Review of Financial Studies 21, 785–818.

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

Barber, B.M., Odean, T., Zhu N., 2009b. Systematic noise. Journal of Financial Markets 12, 547–569.

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

Barberis N., Shleifer A., 2003. Style investing. Journal of Financial Economics 68, 161–199.

Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. Journal of Financial Economics 75, 283–318.

Chan, L.K.C., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. Journal of Finance 51, 1681–1713.

Chemmanur, T.J., Yan, A., 2009. Advertising, attention, and stock returns. Working Paper.

Chordia, T., Shivakumar, L., 2006. Earnings and price momentum. Journal of Financial Economics 80, 627–656.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. Journal of Finance 66, 1461-1499.

DellaVigna, S., Pollet, J.M., 2009. Investor inattention and Friday earnings announcements. Journal of Finance 64, 709–749.

Durnev, A., Morck, R., Yeung, B., 2004. Value-enhancing capital budgeting and firm-specific stock return variation. Journal of Finance 59, 65–105.

Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. Journal of Finance 64, 2023–2052.

Fang, L.H., Peress, J., Zheng, L., 2014. Does media coverage of stocks affect mutual funds' trading and performance? Review of Financial Studies 27, 3441–3466.

Gao, X., Lin, T.C., 2015. Do individual investors treat trading as a fun and exciting gambling activity? Evidence from repeated natural experiments. Review of Financial Studies 28, 2128–2166.

Gervais, S., Kaniel, R., Mingelgrin, D.H., 2001. The high-volume return premium. Journal of Finance 56, 877–919.

Green, T.C., Hwang, B.-H., 2012. Initial public offerings as lotteries: skewness preference and first-day returns. Management Science 58, 432–444.

Hirshleifer, D., Lim, S.S., Teoh, S.H., 2009. Driven to distraction: extraneous events and underreaction to earnings news. Journal of Finance 64, 2289–2325.

Hou, K., Peng, L., Xiong, W., 2009. A tale of two anomalies: the implications of investor attention for price and earnings momentum. Working Paper.

Kacperczyk, M., Van Nieuwerburgh, S. Veldkamp, L., 2016. A rational theory of mutual funds' attention allocation. Econometrica 84, 571–626.

Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L., 2014. Time-varying fund manager skill. Journal of Finance 69, 1455–1484.

Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Englewood Cliffs, NJ.

Karlsson, N., Loewenstein G., and Seppi D., 2009. The ostrich effect: Selective attention to information. Journal of Risk and Uncertainty 38, 95–115.

Kothari, S.P., 2001. Capital markets research in accounting. Journal of Accounting and Economics 31, 105–231.

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

Kumar, A., Lee, C.M.C., 2006. Retail investor sentiment and return comovements. Journal of Finance 61, 2451-2486.

Kumar A., Page J.K., Spalt O.G., 2013. Investor sentiment and return comovements: Evidence from stock splits and headquarters changes. Review of Finance 17, 921–953.

Kumar A., Page J.K., Spalt O.G., 2016. Gambling and comovement. Journal of Financial and Quantitative Analysis, forthcoming.

Liu, H., Peng, L., 2015. Investor attention: Seasonal patterns and endogenous allocations. Working paper.

Lou, D., 2014. Attracting investor attention through advertising. Review of Financial Studies 27, 1797-1829.

Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: why do emerging markets have synchronous stock price movements?. Journal of Financial Economics 58, 215–260.

Peng, L., 2005. Learning with information capacity constraints. The Journal of Financial and Quantitative Analysis 40, 307–329.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. Journal of Financial Economics 80, 563–602.

Seasholes, M.S., Wu, G., 2007. Predictable behavior, profits, and attention. Journal of Empirical Finance 14, 590–610.

Sicherman, N., Loewenstein G., Seppi D., Utkus S., 2016. Financial Attention. Review of Financial Studies 29, 863-897.

Schmidt, D., 2013. Investors' attention and stock co-variation: Evidence from Google sport searches. Working paper.

Simon, H.A., 1955. A behavioral model of rational choice. Quarterly Journal of Economics 69, 99– 118.

Simon, H.A., 1971. Designing Organizations for an Information-Rich World. The Johns Hopkins Press, Baltimore, MD.

Sims, C.A., 2003. Implications of rational inattention. Journal of Monetary Economics 50, 665–690.

Sims, C.A., 2010. Rational inattention and monetary economics *Handbook of Monetary Economics* 3, 155-181.

Van Nieuwerburgh, S., Veldkamp, L., 2009. Information immobility and the home bias puzzle. Journal of Finance 64, 1187–1215.

Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. Review of Economic Studies 77, 779-805.

Veldkamp L., 2006. Information markets and the comovement of asset prices. Review of Economic Studies 73, 823–845.

Veldkamp, L., Wolfers, J., 2007. Aggregate shocks or aggregate information? Costly information and business cycle Comovement. Journal of Monetary Economics 54, 37–55.

Wurgler, J., 2000. Financial markets and the allocation of capital. Journal of Financial Economics 58, 187–214.

Yuan, Y., 2015. Market-wide attention, trading, and stock returns. Journal of Financial Economics 116, 548–564.

Table 1
Lottery jackpot size descriptive statistics

Year	Number of lottery drawings	Number of drawings (jackpot> 500 mil.)	Distribution of jackpot size (in millions of TWD)					
			Mean	Std.	Min.	50th percentile	90th percentile	Max.
All years	2,654	341	278	309	20	184	575	3,381
2002	99	27	470	238	284	379	698	1,823
2003	104	11	341	182	195	287	522	1,329
2004	209	16	257	238	84	198	430	2,097
2005	208	14	218	220	65	149	431	1,732
2006	248	15	183	283	42	88	378	2,930
2007	209	11	147	202	23	77	369	1,799
2008	209	37	317	361	27	195	770	2,138
2009	209	32	276	216	32	198	599	1,209
2010	209	15	233	181	36	180	457	1,310
2011	208	16	245	256	27	186	470	2,783
2012	209	44	337	342	23	216	777	2,281
2013	212	43	343	400	20	199	815	2,837
2014	212	31	285	299	31	170	663	1,650
2015	109	29	490	629	32	240	1,446	3,381

Reported are the descriptive statistics for the lottery jackpots from January 22, 2002, to June 30, 2015. There are a total of 2,654 lottery drawings, and reported values are in millions of Taiwan dollars (TWD). Column (3) reports the number of lottery drawings with a jackpot larger than 500 million TWD. Displayed in column (4) to (9) are mean, standard deviation (denoted Std.), minimum (denoted Min.), 50th percentile, 90th percentile, and maximum (denoted Max.) of the jackpots.

Table 2
Individual firms' return co-movement with market

		Large jackpot days	Non-large jackpot days	Difference (p value)
(1) Correlation coefficient	Mean	0.463	0.445	0.018*** (<.0001)
	Median	0.489	0.459	0.023*** (<.0001)
(2) Adjusted R-squared	Mean	0.232	0.211	0.021*** (<.0001)
	Median	0.235	0.210	0.018*** (<.0001)
Number of firms		817	817	817

(1) Correlation coefficient: we calculate time-series Pearson correlation of individual stock return and market return (all adjusted by daily risk free rate, which is calculated from rate on central bank one-month time deposit) for large jackpot days and non-large jackpot days separately. For each firm, we obtain the paired difference in correlation calculated from large and non-large jackpot days. Reported are mean and median of correlation coefficient, and mean and median of the change on large jackpot days compared with non-large jackpot days among 817 firms.
(2) Adjusted R-squared: for large jackpot days and non-large jackpot days separately, we run the following regression for firm i to obtain adjusted R-squared:

$$Ret_{i,t} = \alpha_i + \beta_i * MktRet_t + \varepsilon_{i,t}$$

Where $Ret_{i,t}$ is excess return of firm i at time t; $MktRet_t$ stands for market excess return at time t. Then we compute the paired difference of adjusted R-squared between large jackpot days and non-large jackpot days for each firm. Reported are mean and median of adjusted R-squared, and mean and median of the change on large jackpot days compared with non-large jackpot days among 817 firms.
Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 3

Individual firms' return co-movement with industry

		Large jackpot days	Non-large jackpot days	Difference (p value)
(1) Correlation coefficient	Mean	0.512	0.500	0.012*** (<.0001)
	Median	0.531	0.496	0.017*** (<.0001)
(2) Adjusted R-squared	Mean	0.291	0.275	0.016*** (<.0001)
	Median	0.280	0.246	0.015*** (<.0001)
Number of firms		817	817	817

(1) Correlation coefficient: we calculate time-series Pearson correlation of stock return and the corresponding industry return (all adjusted by daily risk free rate, which is calculated from rate on central bank one-month time deposit) for large jackpot days and non-large jackpot days separately. For each firm, we obtain the paired difference in correlation calculated from large and non-large jackpot days. Reported are mean and median of correlation coefficient, and mean and median of the change on large jackpot days compared with non-large jackpot days among 817 firms.

(2) Adjusted R-squared: for large jackpot days and non-large jackpot days separately, we run the following regression for firm i to obtain adjusted R-squared:

$$Ret_{i,t} = \alpha_i + \beta_i * IndRet_{i,t} + \varepsilon_{i,t}$$

Where $Ret_{i,t}$ is excess return of firm i at time t; $IndRet_{i,t}$ stands for excess return of industry that firm i belongs to at time t. Then we compute the paired difference of adjusted R-squared between large jackpot days and non-large jackpot days for each firm. Reported are mean and median of adjusted R-squared, and mean and median of the change on large jackpot days compared with non-large jackpot days among 817 firms.

Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 4

Pooled comparison of individual firm's return co-movement with industry and with market

		Model (1) co-movement with industry	Model (2) co-movement with market	Model difference (2)−(1)
		Difference (p value)	Difference (p value)	Difference (p value)
(1) Correlation coefficient	Mean	0.012***	0.018***	0.005***
		(<.0001)	(<.0001)	(0.001)
	Median	0.017***	0.023***	0.006***
		(<.0001)	(<.0001)	(<.0001)
(1) Adjusted R-squared	Mean	0.016***	0.021***	0.005***
		(<.0001)	(<.0001)	(0.001)
	Median	0.015***	0.018***	0.005***
		(<.0001)	(<.0001)	(0.0001)
Number of firms		817	817	817

Model (1): the procedures are the same with those presented in Table 3;
Model (2): the procedures are the same with those presented in Table 2.
For the last column, we calculate the paired difference of the change in correlation coefficient and adjusted R-squared on large jackpot days between Model (2) and Model (1) for each firm. Reported are the mean and median of increase in correlation coefficient/ adjusted R-squared on large jackpot days for both models, and the difference in that increase between models. Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 5

Portfolio return co-movement with market, portfolios sorted on individual trading ratio

Dependent variable	Coefficient (p value)	$\bar{R}^2$	Number of obs.
(1) Correlation coefficient increase	0.002*** (0.002)	0.321	25
(2) Adjusted R-squared increase	0.002*** (0.006)	0.257	25

Stocks are sorted into 25 portfolios according to average individual trading ratio in past 22 trading days, which is:

$$Itr_{i,t} = \frac{\sum_{t-22}^{t-1} V_{i,t}}{\sum_{t-22}^{t-1} TV_{i,t}}$$

$V_{i,t}$ is individual investor trading volume; $TV_{i,t}$ is total trading volume. Then we calculate equally weighted portfolio returns for each trading day.

(1) Correlation coefficient: we compute time-series Pearson correlation of portfolio excess return and market excess return from large jackpot days and from non-large jackpot days separately, and the change in correlation on large jackpot days. Then we regress the change in correlation on large jackpots on the ranking of individual trading ratio. The coefficient on ranking, p value and $\bar{R}^2$ are reported.

(2) Adjusted R-squared: for each portfolio, we run the following regression on large jackpot days and non-large jackpot days:

$$PortRet_{j,t} = \alpha_j + \beta_j MktRet_t + \varepsilon_{i,t}$$

Where $PortRet_{j,t}$ is excess return of portfolio j at time t and $MktRet_t$ is market excess return at time t. Then we regress the increase of adjusted R-squared on large jackpots on the ranking of individual trading ratio. The coefficient on ranking, p value and $\bar{R}^2$ are reported.

***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 6

Large jackpots and market reactions to earnings announcements

Panel A: Stock returns underreacting to earnings announcements

	Dependent variable: announcement day abnormal return		Dependent variable: CAR(0,1)	
	(1)	(2)	(3)	(4)
Constant	-0.501	0.807***	-1.311***	0.591
	(-8.170)	(2.733)	(-12.750)	(0.806)
Large jackpot	0.246	0.115	0.129	-0.069
	(1.603)	(0.703)	(0.350)	(-0.183)
SUE decile rank	0.080***	0.072***	0.211***	0.207***
	(12.027)	(11.264)	(15.295)	(15.462)
(Large Jackpot)*(SUE decile rank)	-0.056***	-0.047***	-0.096**	-0.089*
	(-4.082)	(-3.298)	(-2.216)	(-1.959)
Standard Controls	No	Yes	No	Yes
R2	0.009	0.046	0.028	0.061
N	21,685	21,492	21,685	21,492

Panel B: Stock returns overreacting after earnings announcement

	Dependent variable: CAR(2,5)		Dependent variable: CAR(6,60)	
	(1)	(2)	(3)	(4)
Constant	-1.199***	-2.242***	-0.813	-13.847***
	(-7.012)	(-2.772)	(-0.835)	(-2.947)
Large jackpot	-1.305**	-1.292*	0.019	1.010
	(-1.998)	(-1.821)	-0.012	-0.908
SUE decile rank	0.193***	0.197***	0.458***	0.496***
	(10.794)	(10.733)	(4.804)	(5.349)
(Large jackpot)*(SUE decile rank)	0.157**	0.137*	-0.067	-0.093
	(2.089)	(1.806)	(-0.458)	(-0.653)
Standard Controls	No	Yes	No	Yes
R2	0.013	0.038	0.007	0.073
N	21,685	21,492	21,685	21,404

In event time, day 0 is the day of the announcement. The cumulative abnormal return is the raw buy-and-hold return adjusted by estimated beta from the market model from 300 trading days to 46 trading days before the date of the announcement. The abnormal return is calculated as $Ret_{i,t} - \overline{\beta}_{i,t} MktRet_t$. $Ret_{i,t}$ is firm excess return; $MktRet_t$ is market excess return. The cumulative abnormal return over the window of $\tau + h \; to \; \tau + H$ is

$$CAR(\tau + h, \tau + H)_{i,t} = [\prod_{j=\tau+h}^{\tau+H} (1 + Ret_{i,j})] - 1 - \overline{\beta}_{i,t}[\prod_{j=\tau+h}^{\tau+H} \left(1 + MktRet_j\right) - 1]$$

Earnings surprise is measured by standardized unexpected earnings (SUE). SUE is calculated as $(eps_q - eps_{q-4})/(\sigma_q)$, where eps_q is the most recently announced earnings, eps_{q-4} is earnings in the same quarter of the previous year, and σ_q is the standard deviation of the difference over the prior eight quarters. SUE are winsorized at the 1st and 99th percentiles. The breakpoints for SUE deciles are decided separately for each quarter (smallest SUE group: SUE decile rank=1; largest SUE group: SUE decile rank=10). Standard controls include industry dummies, decile of market capitalization at announcement day, number of announcements at announcement day, past 22 days average individual trading ratio before announcement day, past 22 days average return before announcement day, past 22 days idiosyncratic volatility before announcement day, year, month, day of the week, turnover at announcement day. Standard errors are adjusted for heteroskedasticity and clustered by the date of announcement and firm. Panel A reports the regression results where dependent variables are announcement day abnormal return and CAR(0,1); Panel B reports the regression results where dependent variables are CAR(2,5) and CAR(6,60). We control for number of large jackpots in (M, N)/(N-M+1) and interact it with SUE decile rank for post announcement analysis. T statistics are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 7

Individual firm's return co-movement with industry and with market: by Google search volume index

		Model (1) co-movement with industry			Model (2) co-movement with market			Model difference (2)−(1)
		Large search volume days	Non-large search volume days	Difference (p value)	Large search volume days	Non-large search volume days	Difference (p value)	Difference (p value)
(1) Correlation coefficient	Mean	0.515	0.495	0.020*** (<.0001)	0.474	0.448	0.026*** (<.0001)	0.005*** (0.0031)
	Median	0.516	0.493	0.020*** (<.0001)	0.493	0.462	0.030*** (<.0001)	0.008*** (<.0001)
(2) Adjusted R-squared	Mean	0.292	0.272	0.020*** (<.0001)	0.238	0.215	0.023*** (<.0001)	0.003* (0.077)
	Median	0.263	0.243	0.018*** (<.0001)	0.240	0.214	0.025*** (<.0001)	0.005*** (0.001)
Number of firms		817	817		817	817		817

The sample period is from January 4, 2004 to June 30, 2015. We manually collect Google weekly search index over the sample period and unadjusted Google daily search index over each quarter. The search word is "樂透" which means "lottery" or "lotto" in English and search region is Taiwan. Since unadjusted daily search index is the relative search volume over each quarter, we calculate adjusted daily search index as follows: Adjusted daily search index= (Weekly search index to which the day belongs)*(Unadjusted daily search index/ Weekly average of unadjusted daily search index). We add one to each unadjusted raw daily search index to more precisely compare the numbers over the sample period. Large Google search volume days are those when the daily search index is bigger than 12.47, which is approximately 90th percentile value of the total sample.
Model (1): Correlation coefficient: for large search volume days and non-large search volume days separately, we calculate Pearson correlation of individual stock return and corresponding industry return (all adjusted by daily risk free rate, which is calculated from rate on central bank one-month time deposit). We obtain correlation coefficients for firms for both large and non-large search volume days. Then we compare the paired mean/median difference of correlation coefficients between large and non-large search volume days among all firms.
Adjusted R-squared: for large search volume days and non-large search volume days separately, we run the following regression for firm i:

$$Ret_{i,t} = \alpha_i + \beta_i IndRet_{i,t} + \varepsilon_{i,t}$$

Where $Ret_{i,t}$ represents excess return of firm i at time t; $IndRet_{i,t}$ stands for industry excess return at time t.

We obtain adjusted R-squared for firms, during large search volume days and non-large search volume days. Then we compare the paired mean/median difference of adjusted R-squared between the two groups among all firms.

Model (2): procedures are the same with model (1) except that we replace industry excess return with market excess return. Reported are the increase in correlation coefficient/adjusted R-squared on large search volume days for both models, and the difference in that increase between models. Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 8

Portfolio return co-movement with market, portfolios sorted on retail investor preference net of lottery-type features

Dependent variable	Coefficient (p value)	$\bar{R}^2$	Number of obs.
(1) Correlation coefficient increase	0.001** (0.042)	0.131	25
(2) Adjusted R-squared increase	0.001* (0.077)	0.092	25

We firstly calculate individual trading ratio in the past 22 trading days, average closing price in the past 22 trading days, idiosyncratic skewness and idiosyncratic volatility in the past 150 trading days for each firm-date observation. Each date, we perform cross-sectional regressions of individual trading ratio (past 22 trading days) on average closing price (past 22 trading days), idiosyncratic skewness and idiosyncratic volatility (past 150 trading days), obtaining residuals, which represent characteristics prefered by retail investors other than lottery-type features. Then we divide the ordered distribution of stocks by the residual into 25 parts, each respectively representing the bottom 4% to the top 4% of the sample and calculate the equally weighted portfolio excess returns.

(1) Correlation coefficient: we compute time-series Pearson correlation of portfolio excess return and market excess return from large jackpot days and from non-large jackpot days separately, and the change in correlation on large jackpot days. Then we regress the change in correlation on large jackpots on the ranking of individual trading ratio. The coefficient on ranking, p value and $\bar{R}^2$ are reported.

(2) Adjusted R-squared: for each portfolio, we run the following regression on large jackpot days and non-large jackpot days:

$$PortRet_{j,t} = \alpha_j + \beta_j MktRet_t + \varepsilon_{i,t}$$

Where $PortRet_{j,t}$ is excess return of portfolio j at time t and $MktRet_t$ is market excess return at time t. Then we regress the increase of adjusted R-squared on large jackpots on the ranking of individual trading ratio. The coefficient on ranking, p value and $\bar{R}^2$ are reported.
***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 9

Size of the lottery jackpot and individual firm return co-movement with market

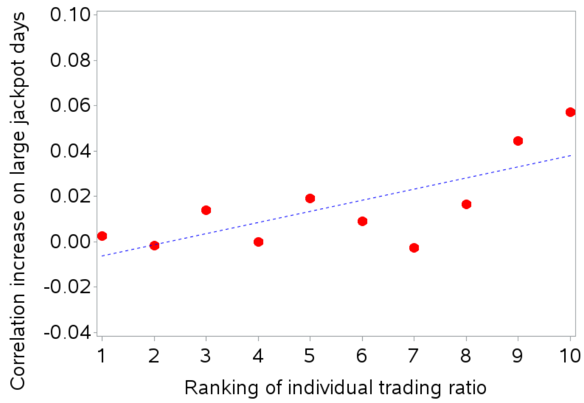
		Cutoff: 400 mil. TWD			Cutoff: 500 mil. TWD			Cutoff: 600 mil. TWD		
		Large jackpot days	Non-large jackpot days	Difference (p value)	Large jackpot days	Non-large jackpot days	Difference (p value)	Large jackpot days	Non-large jackpot days	Difference (p value)
(1) Correlation coefficient	Mean	0.454	0.446	0.008*** (0.002)	0.463	0.445	0.018*** (<.0001)	0.481	0.444	0.038*** (<.0001)
	Median	0.475	0.458	0.014*** (<.0001)	0.489	0.459	0.023*** (<.0001)	0.513	0.456	0.049*** (<.0001)
(2) Adjusted R-squared	Mean	0.222	0.212	0.011*** (<.0001)	0.232	0.211	0.021*** (<.0001)	0.253	0.210	0.043*** (<.0001)
	Median	0.224	0.210	0.010*** (<.0001)	0.235	0.210	0.018*** (<.0001)	0.259	0.208	0.042*** (<.0001)
Number of firms		817	817		817	817		817	817	

We compare the incorporation of market information in stock prices on large jackpot days with different cutoffs. The cutoff varies from 400 million TWD to 600 million TWD. We consider those days with jackpot bigger than the cutoff as large jackpot days. For large jackpot days and non-large jackpot days, we calculate time-series Pearson correlation for all stocks, and then compare the mean and median of the correlation on large jackpot days and non-large jackpot days. We also run the following time-series regression for all stocks on large jackpot days and non-large jackpot days separately:

$$Ret_{i,t} = \alpha_i + \beta_i MktRet_t + \varepsilon_{i,t}$$

Where $Ret_{i,}$ is excess return of firm i at time t and $MktRet_t$ is market excess return at time t. Then the mean and median of the adjusted R-squared for large jackpots, non-large jackpots and their difference are reported. Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Panel A: Portfolio return co-movement with market return



Panel B: Adjusted R-squared from regressing portfolio return on market return

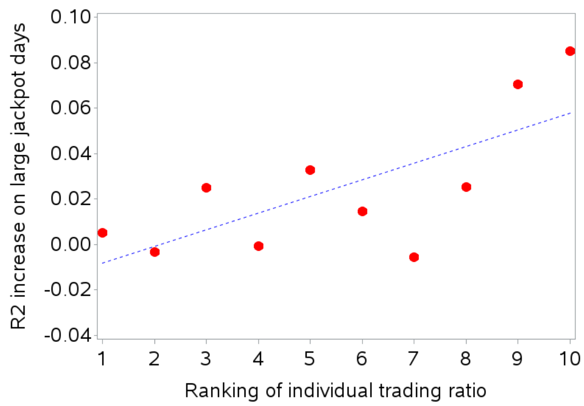


Figure 1. **Impact of large jackpots on portfolio return co-movement with market, portfolios sorted on individual trading ratio/fraction**

Stocks are sorted into 10 portfolios according to average individual trading ratio in past 22 trading days (Itr) and corresponding equally-weighted portfolio returns each trading day are computed.

$$Itr_{i,t} = \frac{\sum_{t-22}^{t-1} V_{i,t}}{\sum_{t-22}^{t-1} TV_{i,t}}$$

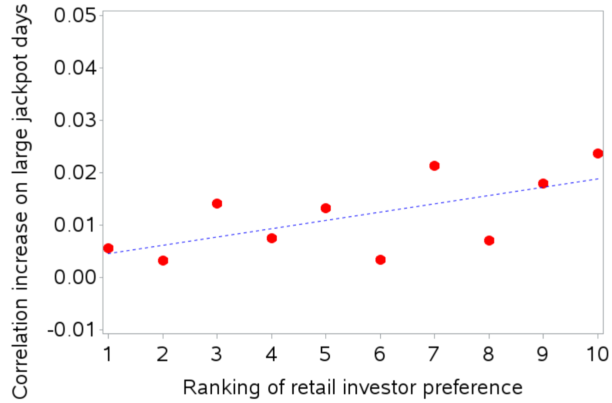
Panel A: we calculate Pearson correlation between excess return of each portfolio and market excess return on large jackpot days and on non-large jackpot days separately, and obtain the difference. The vertical axis reports the increase of correlation coefficients on large jackpot days compared with those on non-large jackpots; horizontal axis reports the ranking of average individual trading ratio.

Panel B: for each portfolio, we run the following time-series OLS regression on large jackpot days and non-large jackpot days:

$$PortRet_{j,t} = \alpha_j + \beta_j MktRet_t + \varepsilon_{j,t}$$

where $PortRet_{j,t}$ is portfolio excess return at time t and $MktRet$ is market excess return at time t. The vertical axis reports the difference of adjusted R-squared on large jackpots and non-large jackpots; horizontal axis reports the ranking of average individual trading ratio.

Panel A: Portfolio return co-movement with market return



Panel B: Adjusted R-squared from regressing portfolio return on market return

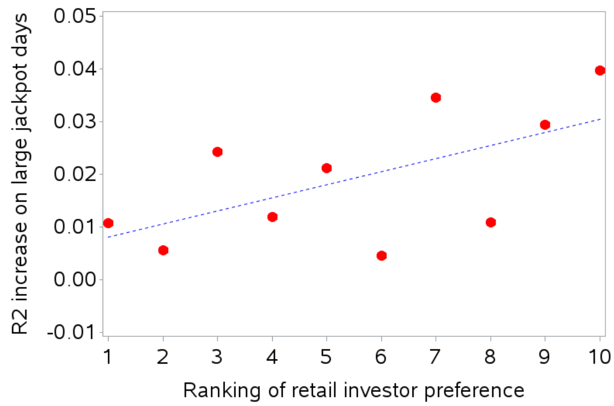


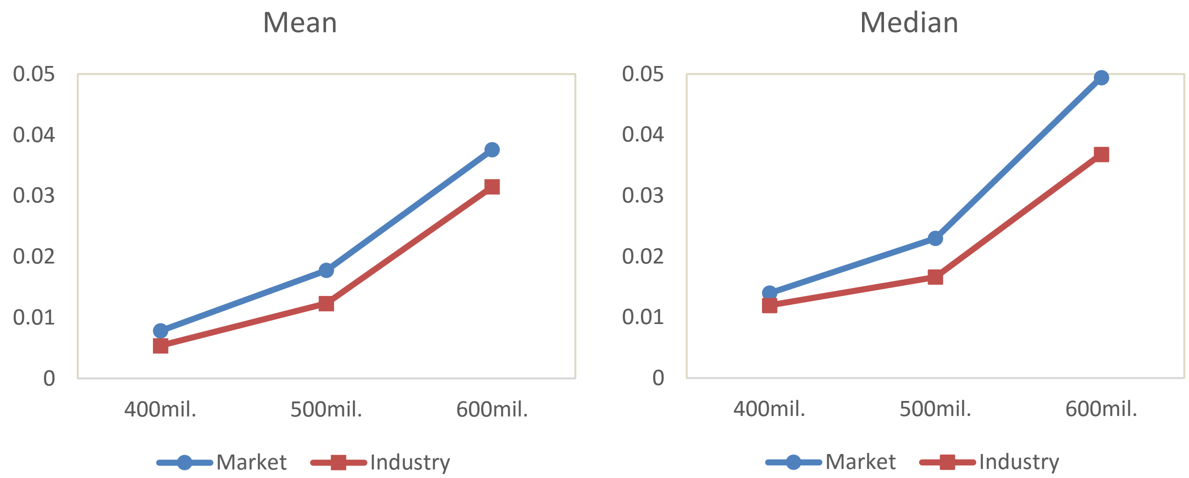
Figure 2. **Impact of large jackpots on portfolio return co-movement with market, portfolios sorted on retail investor preference net of lottery-type features**

This figure shows that stocks more preferred by retail investors co-move more with market, excluding the effect of lottery-type features. We firstly calculate individual trading ratio in the past 22 trading days, average closing price in the past 22 trading days, idiosyncratic skewness and idiosyncratic volatility in the past 150 trading days for each firm-date observation. Each date, we perform cross-sectional regressions of individual trading ratio (past 22 trading days) on average closing price (past 22 trading days), idiosyncratic skewness and idiosyncratic volatility (past 150 trading days), obtaining residuals, which represent characteristics preferred by retail investors other than lottery-type features. Then we divide the ordered distribution of stocks by the residual into 10 parts, each respectively representing the bottom 10% to the top 10% of the sample and calculate the equally weighted portfolio excess returns.
Panel A: we calculate Pearson correlation between excess return of each portfolio and market excess return on large jackpot days and on non-large jackpot days separately, and obtain the difference. The vertical axis reports the increase of correlation coefficients on large jackpot days compared with those on non-large jackpots; horizontal axis reports the ranking of average individual trading ratio.
Panel B: for each portfolio, we run the following time-series OLS regression on large jackpot days and non-large jackpot days:

$$PortRet_{j,t} = \alpha_j + \beta_j MktRet_t + \varepsilon_{j,t}$$

where $PortRet_{j,t}$ is portfolio excess return at time t and $MktRet_t$ is market excess return at time t. The vertical axis reports the difference of adjusted R-squared on large jackpots and non-large jackpots; horizontal axis reports the ranking of average individual trading ratio.

Panel A: Increase in individual firm return correlation with market/industry on large jackpots



Panel B: Increase in adjusted R-squared of regressing individual firm return on market/industry on large jackpots

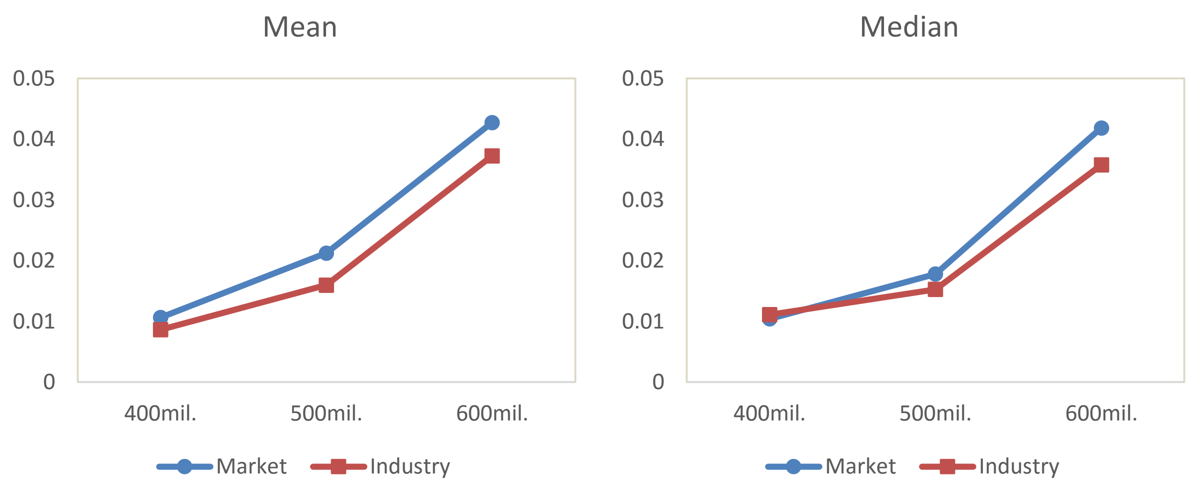
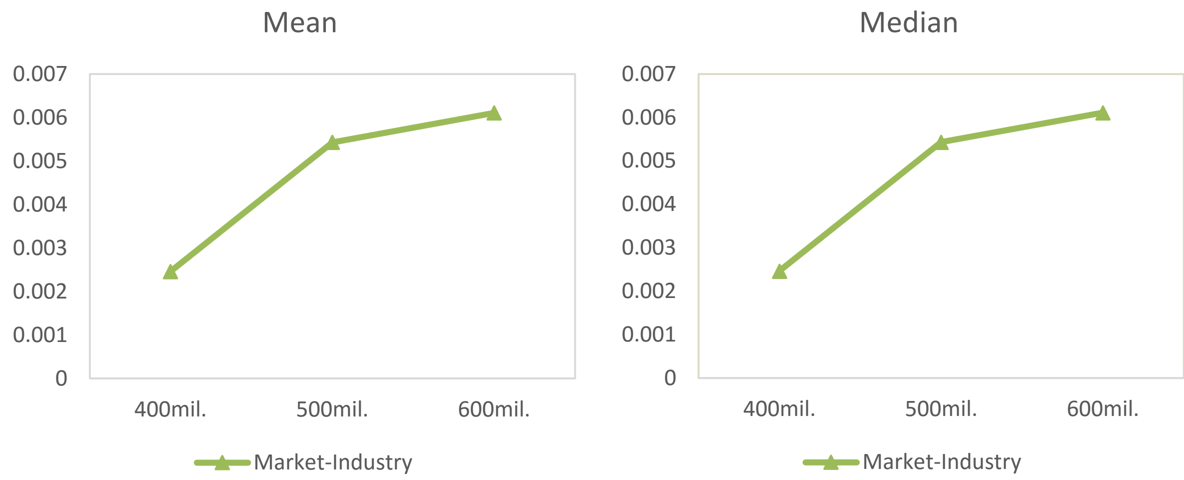


Figure 3. **Impact of size of lottery jackpot on return co-movement with market and with industry**

Plotted is the increase in co-movement with market (circles) and with industry (diamonds) on large jackpot days versus the size of jackpot (on the x-axis). The co-movement is measured by return correlation and adjusted R-squared. The size of the jackpot varies from 400 million TWD to 600 million TWD. In panel A, we calculate the increase in time-series Pearson correlation coefficient of individual firm return and market/industry return during large jackpot days compared with non-large jackpot days. Reported are the mean and median of the increase on large jackpots. In panel B, we calculate the difference of adjusted R-squared in regressing individual firm excess return on market or industry excess return on large jackpot days and on non-large jackpot days. Reported are the mean and median of increase on large jackpots.

Panel A: Increase in individual firm return correlation with market/industry on large jackpots



Panel B: Increase in adjusted R-squared of regressing individual firm return on market/industry on large jackpots

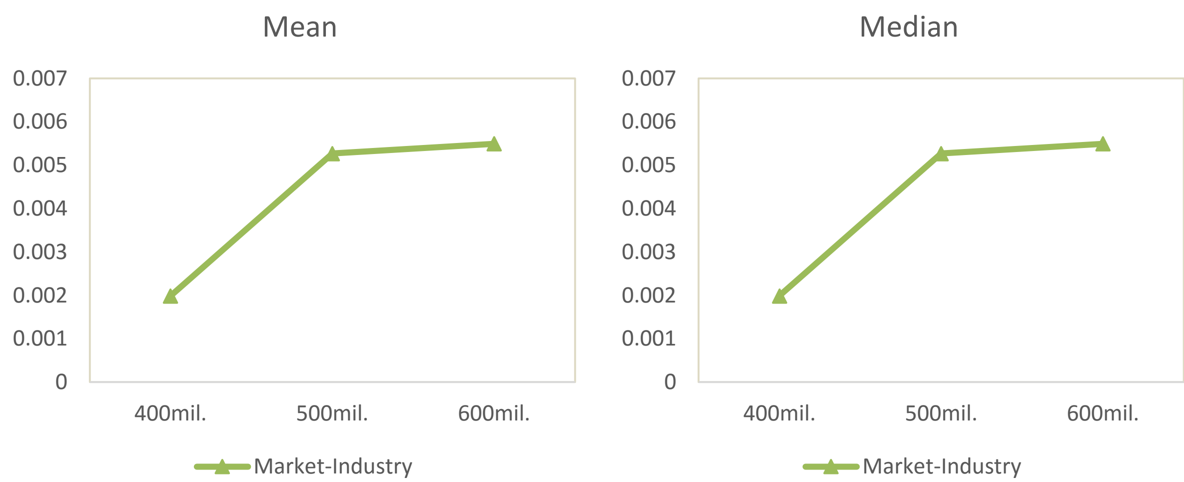


Figure 4. **Impact of size of lottery jackpot on return co-movement: comparison between market and industry**

Plotted is the difference between extra incorporation of market information and of industry (triangles) information on large jackpot days versus the size of jackpot (on the x-axis). The difference is measured by the difference in increase of return correlation and adjusted R-squared on large jackpot days from using market return and industry return. The size of the jackpot varies from 400 million TWD to 600 million TWD. In panel A, we use time-series Pearson correlation coefficient of individual firm return and market/industry return. In panel B, we use adjusted R-squared from regressing individual firm excess return on market or on industry excess return.

Appendix

Table A1

Correlation between lottery jackpot and Google search volume for "lottery"

Panel A: Effect of large jackpot on daily Google search volume

	Large jackpot days	Non-large jackpot days	Difference (p value)
Mean	19.787	7.825	11.961*** (<.0001)
Median	10.606	6.507	4.099*** (<.0001)
N. of days	303	3898	

Panel B: Correlation coefficient

	Tetrachoric correlation coefficient (p value)	Pearson correlation coefficient (p value)
Large jackpot indicator v.s. large Google search volume indicator	0.584*** (<.0001)	
Jackpot size v.s. Google search volume		0.390*** (<.0001)

The sample period is from January 4, 2004 to June 30, 2015. We manually collect Google weekly search index over the sample period and unadjusted Google daily search index over each quarter. The search word is “樂透” which means “lottery” or “lotto” in English and search region is Taiwan. Since unadjusted daily search index is the relative search volume over each quarter, we calculate adjusted daily search index as follows: Adjusted daily search index= (Weekly search index to which the day belongs)*(Unadjusted daily search index/ Weekly average of unadjusted daily search index). We add one to each unadjusted raw daily search index to more precisely compare the numbers over the sample period. In panel A, we compare the mean and median of daily search volume index between large jackpot days and non-large jackpot days. T test is used for testing mean difference and Mood median test is adopted for testing median difference. In panel B, we calculate correlation coefficients between lottery jackpot and Google search volume index. An indicator variable of large Google search volume equals one if the daily search volume index is bigger than 12.47, which is approximately 90th percentile value of the total sample. Reported are Tetrachoric correlation coefficient for testing the correlation between large jackpot indicator and large search volume indicator, and Pearson correlation coefficient for testing correlation between cumulative prize of the lottery and daily search volume index. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table A2

Annual frequency of firms across industries

TSE industry classification	Number of firms 2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
All industries	609	644	662	677	690	710	722	745	764	799	817	817	817	817
Cement	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Food	20	21	21	21	21	21	21	21	21	21	21	21	21	21
Plastic	20	20	20	20	20	20	20	20	20	21	22	22	22	22
Textile	44	45	45	45	46	46	46	46	46	46	46	46	46	46
Machinery	30	31	33	34	35	36	37	39	40	40	41	41	41	41
Wire & cable	13	13	13	13	13	13	13	13	14	14	14	14	14	14
Chemical	34	35	36	36	36	37	38	41	43	47	50	50	50	50
Chemical						24	24	26	26	26	26	26	26	26
Biotech.						12	13	14	16	20	23	23	23	23
Glass & ceramic	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Paper	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Steel	27	27	27	27	27	27	29	29	29	29	29	29	29	29
Rubber	9	9	9	9	10	10	10	10	10	10	10	10	10	10
Auto	4	4	5	5	5	5	5	7	7	7	8	8	8	8
Electronics	240	266	279	291	301	318	325	341	352	374	379	379	379	379
Semiconductor						52	55	57	60	63	64	64	64	64
Computer						52	52	52	53	54	55	55	55	55
Optoelectronics						46	49	53	57	65	67	67	67	67
Communications						31	31	32	34	36	36	36	36	36
Electronic components						74	75	80	82	87	87	87	87	87
Electronic channel						20	20	20	20	20	20	20	20	20

Information						11	11	11	11	12	12	12	12	12
Other electronics						31	31	32	32	33	34	34	34	34
Construction	43	44	45	46	46	46	46	46	48	48	49	49	49	49
Transportation	15	15	15	15	15	15	16	16	17	18	18	18	18	18
Tourism	8	9	9	9	9	9	9	9	10	10	12	12	12	12
Financials	32	33	33	33	33	33	33	33	33	34	35	35	35	35
Retail	12	12	12	12	12	12	12	12	12	13	15	15	15	15
Gas & oil	7	8	8	8	8	8	8	8	8	8	8	8	8	8
Others	33	34	34	35	35	36	36	36	36	41	42	42	42	42

Reported are number of firms in our sample from January 22, 2002, to June 30, 2015 across industries. Before July 2007, industries are classified into 20 groups. After July, 2, 2007, chemical industry is sub-divided into 2 groups: chemical and biotechnology industry; and electronics industry is sub-divided into 8 groups. The corresponding index returns are provided by Taiwan Stock Exchange.

Table A3
Individual firms' return co-movement with industry

TSE industry classification	Mean			Median		
	Large jackpot days	Non-large jackpot days	Difference (p value)	Large jackpot days	Non-large jackpot days	Difference (p value)
Panel A: Correlation coefficient						
Cement	0.703	0.712	-0.009	0.681	0.616	-0.023
			(0.606)			(0.297)
Food	0.566	0.514	0.052***	0.585	0.525	0.059***
			(<.0001)			(0.0002)
Plastic	0.459	0.452	0.007	0.471	0.437	0.016
			(0.433)			(0.302)
Textile	0.465	0.455	0.011	0.488	0.462	0.015
			(0.227)			(0.144)
Machinery	0.481	0.474	0.006	0.502	0.484	0.011
			(0.582)			(0.315)
Wire & cable	0.595	0.574	0.02	0.576	0.551	0.02
			(0.166)			(0.135)
Chemical	0.485	0.498	-0.013	0.497	0.490	0.007
			(0.437)			(0.883)
Chemical	0.521	0.513	0.007	0.552	0.540	0.023*
			(0.732)			(0.082)
Biotech.	0.589	0.631	-0.042	0.620	0.678	-0.019*
			(0.182)			(0.095)
Glass & ceramic	0.661	0.651	0.01	0.580	0.553	0.022
			(0.723)			(0.875)
Paper	0.758	0.762	-0.004	0.797	0.779	-0.002
			(0.783)			(0.813)

Steel	0.556	0.568	-0.012 (0.337)	0.587	0.605	-0.009 (0.237)
Rubber	0.607	0.612	-0.004 (0.849)	0.604	0.624	0.011 (0.846)
Auto	0.525	0.519	0.007 (0.77)	0.673	0.606	0.003 (0.844)
Electronics	0.489	0.471	0.018*** (<.0001)	0.515	0.478	0.024*** (<.0001)
Semiconductor	0.509	0.460	0.05*** (<.0001)	0.553	0.473	0.054*** (<.0001)
Computer	0.524	0.471	0.052*** (<.0001)	0.571	0.470	0.063*** (<.0001)
Optoelectronics	0.516	0.506	0.01 (0.413)	0.523	0.511	0.024** (0.026)
Communications	0.388	0.384	0.004 (0.749)	0.400	0.386	0.012 (0.378)
Electronic components	0.550	0.506	0.044*** (<.0001)	0.575	0.537	0.045*** (<.0001)
Electronic channel	0.640	0.562	0.078*** (<.0001)	0.654	0.507	0.062*** (<.0001)
Information	0.527	0.545	-0.018 (0.496)	0.523	0.524	-0.001 (0.622)
Other electronics	0.467	0.393	0.074*** (<.0001)	0.485	0.391	0.091*** (<.0001)
Construction	0.524	0.526	-0.002 (0.826)	0.612	0.618	-0.007 (0.965)
Transportation	0.646	0.609	0.037** (0.017)	0.680	0.655	0.031** (0.024)
Tourism	0.616	0.542	0.074* (0.061)	0.663	0.613	0.036** (0.034)
Financials	0.708	0.696	0.012 (0.277)	0.743	0.725	0.007 (0.309)

Retail	0.451	0.438	0.014 (0.386)	0.440	0.412	0.021 (0.389)
Gas & oil	0.371	0.366	0.005 (0.826)	0.293	0.281	0.013 (0.547)
Others	0.445	0.448	-0.003 (0.848)	0.454	0.446	0.01 (0.663)

Panel B: Adjusted R-squared

Cement	0.511	0.525	-0.014 (0.526)	0.462	0.379	-0.030 (0.297)
Food	0.347	0.288	0.059*** (<.0001)	0.340	0.275	0.064*** (<.0001)
Plastic	0.260	0.249	0.011 (0.114)	0.220	0.194	0.011 (0.195)
Textile	0.243	0.231	0.012 (0.126)	0.235	0.213	0.013 (0.132)
Machinery	0.247	0.235	0.012 (0.234)	0.249	0.234	0.006 (0.273)
Wire & cable	0.368	0.348	0.020 (0.158)	0.330	0.304	0.021 (0.135)
Chemical	0.254	0.260	-0.005 (0.686)	0.244	0.240	0.004 (0.823)
Chemical	0.303	0.281	0.022 (0.21)	0.302	0.291	0.024** (0.044)
Biotech.	0.373	0.416	-0.043 (0.134)	0.380	0.459	-0.033 (0.101)
Glass & ceramic	0.472	0.460	0.012 (0.705)	0.334	0.305	0.027 (0.875)
Paper	0.582	0.588	-0.006 (0.759)	0.634	0.607	-0.004 (0.813)

Steel	0.331	0.345	-0.014 (0.217)	0.342	0.365	-0.011 (0.204)
Rubber	0.400	0.400	0.001 (0.979)	0.365	0.389	0.002 (1)
Auto	0.353	0.340	0.013 (0.521)	0.452	0.372	0.004 (0.844)
Electronics	0.254	0.233	0.021*** (<.0001)	0.263	0.228	0.021*** (<.0001)
Semiconductor	0.282	0.228	0.055*** (<.0001)	0.303	0.223	0.054*** (<.0001)
Computer	0.299	0.240	0.059*** (<.0001)	0.323	0.221	0.058*** (<.0001)
Optoelectronics	0.288	0.271	0.017* (0.075)	0.269	0.261	0.02** (0.027)
Communications	0.165	0.160	0.005 (0.523)	0.156	0.148	0.006 (0.378)
Electronic components	0.318	0.271	0.047*** (<.0001)	0.327	0.288	0.043*** (<.0001)
Electronic channel	0.424	0.329	0.095*** (<.0001)	0.425	0.256	0.077*** (<.0001)
Information	0.291	0.306	-0.015 (0.545)	0.271	0.275	-0.004 (0.519)
Other electronics	0.246	0.175	0.071*** (<.0001)	0.232	0.153	0.079*** (<.0001)
Construction	0.331	0.330	0.000 (0.989)	0.372	0.382	-0.004 (0.833)
Transportation	0.429	0.394	0.035** (0.032)	0.461	0.429	0.039** (0.043)
Tourism	0.419	0.345	0.074** (0.019)	0.439	0.381	0.053*** (0.009)
Financials	0.512	0.499	0.014 (0.259)	0.550	0.554	0.010 (0.341)

Retail	0.234	0.217	0.017	0.191	0.525	0.020
			(0.200)			(0.229)
Gas & oil	0.192	0.190	0.002	0.082	0.078	0.008
			(0.867)			(0.641)
Others	0.220	0.217	0.003	0.204	0.199	0.007
			(0.762)			(0.717)

We firstly divide all stocks into 30 industry subsamples. For each subsample, the analyses are as follows:
Panel A: Correlation coefficient: we calculate time-series Pearson correlation of individual stock return and the corresponding industry return (all adjusted by daily risk free rate, which is calculated from rate on central bank one-month time deposit) for large jackpot days and non-large jackpot days separately. For each firm, we obtain the paired difference in correlation calculated from large and non-large jackpot days. Reported are mean and median of correlation coefficient, and mean and median of the change on large jackpot days compared with non-large jackpot days among firms in each industry.
Panel B: Adjusted R-squared: for large jackpot days and non-large jackpot days separately, we run the following regression for firm i:

$$Ret_{i,t} = \alpha_i + \beta_i IndRet_{i,t} + \varepsilon_{i,t}$$

Where $Ret_{i,t}$ is excess return of firm i at time t; $IndRet_{i,t}$ stands for excess return of industry that firm i belongs to at time t. Then we compute the paired difference of adjusted R-squared between the two groups. Reported are mean and median of adjusted R-squared, and mean and median of the change on large jackpot days compared with non-large jackpot days among firms in each industry.
Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table A4

The comparison between individual firm return co-movement with industry and with market

TSE industry classification	Model (1) co-movement with industry		Model (2) co-movement with market		Model difference (2)−(1)	
	Mean difference (p value)	Median difference (p value)	Mean difference (p value)	Median difference (p value)	Mean difference (p value)	Median difference (p value)
Panel A: Correlation coefficient						
Cement	-0.009	-0.023	-0.002	0.014	0.007	0.021
	(0.606)	(0.297)	(0.909)	(0.938)	(0.659)	(0.688)
Food	0.052***	0.059***	0.074***	0.083***	0.022**	0.03**
	(<.0001)	(0.00)	(<.0001)	(<.0001)	(0.018)	(0.017)
Plastic	0.007	0.016	-0.017	-0.006	-0.025**	-0.018*
	(0.433)	(0.302)	(0.15)	(0.184)	(0.044)	(0.051)
Textile	0.011	0.015	-0.009	0.001	-0.02***	-0.02***
	(0.227)	(0.144)	(0.325)	(0.617)	(0.0003)	(0.0001)
Machinery	0.006	0.011	0.018	0.018	0.011	0.014
	(0.582)	(0.315)	(0.204)	(0.117)	(0.226)	(0.11)
Wire & cable	0.02	0.02	0.023	0.027	0.002	0.007
	(0.166)	(0.135)	(0.132)	(0.153)	(0.792)	(0.542)
Chemical	-0.013	0.007	0.001	0.016	0.014	0.01*
	(0.437)	(0.883)	(0.931)	(0.137)	(0.119)	(0.061)
Chemical	0.007	0.023*	0.015	0.041**	0.008	0.005
	(0.732)	(0.082)	(0.503)	(0.029)	(0.483)	(0.267)
Biotech.	-0.042	-0.019*	-0.015	0.008	0.026	0.043
	(0.182)	(0.095)	(0.554)	(0.906)	(0.191)	(0.116)
Glass & ceramic	0.01	0.022	0.016	0.021	0.006	0.003
	(0.723)	(0.875)	(0.609)	(0.625)	(0.82)	(0.875)
Paper	-0.004	-0.002	0.042	0.049	0.046**	0.044**

	(0.783)	(0.813)	(0.108)	(0.219)	(0.011)	(0.016)
Steel	-0.012	-0.009	0.018*	0.027***	0.03***	0.037***
	(0.337)	(0.237)	(0.065)	(0.002)	(0.003)	(0.002)
Rubber	-0.004	0.011	-0.009	-0.027	-0.005	-0.013
	(0.849)	(0.846)	(0.696)	(0.557)	(0.812)	(0.846)
Auto	0.007	0.003	0.008	0.019	0.002	0.032
	(0.77)	(0.844)	(0.701)	(0.641)	(0.953)	(0.844)
Electronics	0.018***	0.024***	0.021***	0.027***	0.003***	0.005***
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	0.006	(0.0004)
Semiconductor	0.05***	0.054***	0.042***	0.059***	-0.008*	-0.008*
	(<.0001)	(<.0001)	(0.0004)	(0.0001)	(0.066)	(0.099)
Computer	0.052***	0.063***	0.051***	0.068***	-0.001	-0.003
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.7)	(0.579)
Optoelectronics	0.01	0.024**	0.018	0.032***	0.008	0.009*
	(0.413)	(0.026)	(0.109)	(0.0049)	(0.253)	(0.082)
Communications	0.004	0.012	0.031**	0.043**	0.027***	0.031***
	(0.749)	(0.378)	(0.02)	(0.012)	(0.0005)	(0.0005)
Electronic components	0.044***	0.045***	0.062***	0.067***	0.018***	0.02***
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Electronic channel	0.078***	0.062***	0.071***	0.072***	-0.007	-0.007
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.504)	(0.475)
Information	-0.018	-0.001	0.014	0.040	0.031***	0.034***
	(0.496)	(0.622)	(0.561)	(0.4238)	(0.006)	(0.009)
Other electronics	0.074***	0.091***	0.058***	0.069***	-0.016*	-0.023**
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.051)	(0.021)
Construction	-0.002	-0.007	0.003	0.008	0.005	0.016
	(0.826)	(0.965)	(0.768)	(0.464)	(0.506)	(0.187)
Transportation	0.037**	0.031**	0.041*	0.046*	0.003	0.016
	(0.017)	(0.024)	(0.079)	(0.067)	(0.842)	(0.468)
Tourism	0.074*	0.036**	0.041**	0.042**	-0.033	-0.024
	(0.061)	(0.034)	(0.042)	(0.043)	(0.238)	(0.339)

Financials	0.012	0.007	0.037***	0.037***	0.025***	0.023***
	(0.277)	(0.309)	(0.002)	(0.0005)	(<.0001)	(<.0001)
Retail	0.014	0.021	0.039***	0.04**	0.026**	0.041*
	(0.386)	(0.389)	(0.005)	(0.01)	(0.049)	(0.055)
Gas & oil	0.005	0.013	0.055**	0.051**	0.050	0.032
	(0.826)	(0.547)	(0.013)	(0.016)	(0.113)	(0.109)
Others	-0.003	0.01	0.002	0.013	0.004	0.001
	(0.848)	(0.663)	(0.903)	(0.494)	(0.541)	(0.456)

Panel B: Adjusted R-squared

Cement	-0.014	-0.030	-0.003	0.016	0.011	0.030
	(0.526)	(0.297)	(0.835)	(0.938)	(0.587)	(0.469)
Food	0.059***	0.064***	0.069***	0.071***	0.010	0.021
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.237)	(0.254)
Plastic	0.011	0.011	-0.016	-0.009*	-0.027**	-0.018
	(0.114)	(0.195)	(0.107)	(0.095)	(0.012)	(0.02)
Textile	0.012	0.013	-0.006	-0.001	-0.018***	-0.015***
	(0.126)	(0.132)	(0.392)	(0.655)	(0.001)	(0.0002)
Machinery	0.012	0.006	0.023**	0.017*	0.011	0.011
	(0.234)	(0.273)	(0.034)	(0.055)	(0.195)	(0.12)
Wire & cable	0.020	0.021	0.018	0.028	-0.003	0.002
	(0.158)	(0.135)	(0.171)	(0.173)	(0.785)	(0.903)
Chemical	-0.005	0.004	0.012	0.008	0.017**	0.011**
	(0.686)	(0.823)	(0.261)	(0.107)	(0.024)	(0.027)
Chemical	0.022	0.024**	0.031*	0.037**	0.008	0.005
	(0.21)	(0.044)	(0.056)	(0.027)	(0.293)	(0.364)
Biotech.	-0.043	-0.033	0.000	0.002	0.043**	0.054**
	(0.134)	(0.101)	(0.991)	(0.883)	(0.038)	(0.011)
Glass & ceramic	0.012	0.027	0.019	0.020	0.007	0.004
	(0.705)	(0.875)	(0.593)	(0.625)	(0.815)	(0.875)

Paper	-0.006	-0.004	0.044	0.038	0.05***	0.047**
	(0.759)	(0.813)	(0.118)	(0.156)	(0.003)	(0.016)
Steel	-0.014	-0.011	0.019***	0.019***	0.033***	0.042***
	(0.217)	(0.204)	(0.008)	(0.001)	(0.002)	(0.002)
Rubber	0.001	0.002	-0.004	-0.027	-0.005	0.003
	(0.979)	(1)	(0.841)	(0.375)	(0.821)	(0.922)
Auto	0.013	0.004	0.015	0.016	0.002	0.024
	(0.521)	(0.844)	(0.35)	(0.461)	(0.915)	(0.641)
Electronics	0.021***	0.021***	0.025***	0.022***	0.004***	0.004***
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Semiconductor	0.055***	0.054***	0.053***	0.062***	-0.002	-0.002
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.661)	(0.495)
Computer	0.059***	0.058***	0.058***	0.057***	-0.001	-0.002
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.863)	(0.785)
Optoelectronics	0.017*	0.02**	0.022**	0.029***	0.005	0.010
	(0.075)	(0.027)	(0.014)	(0.004)	(0.377)	(0.105)
Communications	0.005	0.006	0.034***	0.037***	0.028***	0.03***
	(0.523)	(0.378)	(0.007)	(0.005)	(0.0004)	(0.0005)
Electronic components	0.047***	0.043***	0.065***	0.06***	0.018***	0.016***
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Electronic channel	0.095***	0.077***	0.084***	0.067***	-0.012	-0.008
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.332)	(0.388)
Information	-0.015	-0.004	0.018	0.033	0.033**	0.047**
	(0.545)	(0.519)	(0.334)	(0.339)	(0.015)	(0.034)
Other electronics	0.071***	0.079***	0.066***	0.071***	0.066***	0.071***
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)
Construction	0.000	-0.004	0.005	0.002	0.005	0.007
	(0.989)	(0.833)	(0.525)	(0.501)	(0.481)	(0.363)
Transportation	0.035**	0.039**	0.039**	0.048*	0.005	0.003
	(0.032)	(0.043)	(0.042)	(0.054)	(0.751)	(0.702)
Tourism	0.074**	0.053***	0.027*	0.028**	-0.047*	-0.028*

	(0.019)	(0.009)	(0.056)	(0.043)	(0.08)	(0.064)
Financials	0.014	0.010	0.044***	0.048***	0.031***	0.029
	(0.259)	(0.341)	(0.0003)	(0.0001)	(<.0001)	(<.0001)
Retail	0.017	0.020	0.034***	0.032**	0.016	0.022
	(0.200)	(0.229)	(0.009)	(0.013)	(0.135)	(0.121)
Gas & oil	0.002	0.008	0.045**	0.035**	0.043*	0.028
	(0.867)	(0.641)	(0.031)	(0.023)	(0.077)	(0.11)
Others	0.003	0.007	0.011	0.006	0.008*	0.005
	(0.762)	(0.717)	(0.258)	(0.358)	(0.085)	(0.134)

This table reports the comparison of stock return's co-movement with market and with industry. In model (1), we present the results in Table A3 again. In model (2), we repeat the procedures in Table A3 except that industry excess return is replaced with market excess return. Reported are the mean and median of correlation and adjusted R-squared, and the differences in mean and median during large jackpot days and during non-large jackpot days. In last two columns, we calculate the model differences (model (2) minus model (1)) in difference (in mean and median) on large jackpot days and on non-large jackpot days. Reported are mean and median differences and corresponding p value. Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table A5

Pair-wise correlation of individual firm return on large jackpots

TSE industry classification	Mean: Large jackpot days	Mean: Non-large jackpot days	Mean: Difference (p value)	Median: Large jackpot days	Median: Non-large jackpot days	Median: Difference (p value)
All industries	0.268	0.256	0.012***	0.276	0.257	0.02***
			(<.0001)			(<.0001)
Cement	0.557	0.576	-0.018	0.563	0.546	-0.024*
			(-0.132)			(-0.085)
Food	0.424	0.361	0.063***	0.443	0.378	0.063***
			(<.0001)			(<.0001)
Plastic	0.333	0.328	0.006	0.319	0.304	0.011**
			(0.182)			(0.019)
Textile	0.284	0.290	-0.007***	0.279	0.286	-0.005***
			(0.001)			(0.003)
Machinery	0.235	0.237	-0.002	0.234	0.233	0.006
			(0.565)			(0.756)
Wire & cable	0.379	0.365	0.013**	0.365	0.360	0.015**
			(0.049)			(0.018)
Chemical	0.111	0.121	-0.01***	0.088	0.105	-0.018***
			(<.0001)			(<.0001)
Chemical	0.304	0.309	-0.006	0.308	0.306	0.01
			(0.344)			(0.674)
Biotech.	0.309	0.373	-0.064***	0.295	0.364	-0.058***
			(<.0001)			(<.0001)
Glass & ceramic	0.599	0.585	0.014	0.595	0.599	0.008
			(0.455)			(0.563)
Paper	0.564	0.568	-0.004	0.543	0.561	-0.003

			(0.645)			(0.776)
Steel	0.396	0.348	0.048***	0.424	0.357	0.05***
			(<.0001)			(<.0001)
Rubber	0.376	0.399	-0.022*	0.344	0.398	-0.025*
			(0.09)			(0.055)
Auto	0.265	0.271	-0.006	0.207	0.187	0
			(0.677)			(0.677)
Electronics	0.301	0.288	0.013***	0.313	0.297	0.019***
			(<.0001)			(<.0001)
Semiconductor	0.382	0.353	0.029***	0.395	0.355	0.035***
			(<.0001)			(<.0001)
Computer	0.318	0.274	0.044***	0.324	0.287	0.054***
			(<.0001)			(<.0001)
Optoelectronics	0.297	0.308	-0.011***	0.307	0.315	0.004
			(<.0001)			(0.157)
Communications	0.303	0.293	0.01***	0.310	0.309	0.016***
			(<.0001)			(<.0001)
Electronic components	0.354	0.322	0.032***	0.367	0.331	0.039***
			(<.0001)			(<.0001)
Electronic channel	0.470	0.402	0.068***	0.453	0.398	0.064***
			(<.0001)			(<.0001)
Information	0.336	0.342	-0.006	0.337	0.324	-0.004
			(0.455)			(0.537)
Other electronics	0.329	0.282	0.047***	0.336	0.289	0.054***
			(<.0001)			(<.0001)
Construction	0.275	0.282	-0.006***	0.244	0.246	-0.007**
			(0.003)			(0.011)
Transportation	0.415	0.372	0.043***	0.408	0.366	0.035***
			(<.0001)			(<.0001)
Tourism	0.331	0.291	0.039***	0.268	0.225	0.043***
			(0.0003)			(<.0001)
Financials	0.535	0.514	0.021***	0.538	0.526	0.014***

			(<.0001)			(<.0001)
Retail	0.261	0.244	0.017**	0.246	0.224	0.014**
			(0.022)			(0.012)
Gas & oil	0.257	0.251	0.007	0.230	0.224	0.009
			(0.563)			(0.397)
Others	0.201	0.209	-0.008***	0.205	0.205	-0.004
			(0.009)			(0.1)

For firm i and firm j in all industries, i≠j, pair-wise (i, j) time series Pearson correlations are calculated separately on large jackpots and non-large jackpots. We compare the mean and median difference of all pair-wise correlations on large jackpots. The same procedures apply to different industry subsamples. Paired t test is employed for testing mean difference and and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table A6

Portfolios preferred by retail investors and return co-movement with market

	Dependent variable: change in correlation on large jackpot days			Dependent variable: change in adjusted R-squared on large jackpot days		
	Coefficient (p value)	R^2	Number of obs.	Coefficient (p value)	R^2	Number of obs.
(1) Average market cap. (past 22 days)	-0.001***	0.403	25	-0.001***	0.278	25
	(0.001)			(0.007)		
(2) Idiosyncratic volatility (past 180 days)	0.0006**	0.230	25	0.0009**	0.157	25
	(0.015)			(0.050)		

This table shows that portfolios more preferred by retail investors co-move more with market. For daily observations, we sort stocks into 25 portfolios according to (1) each stock's average market capitalization in the past 22 trading days; (2) idiosyncratic return volatility in the past 180 trading days. For each portfolio, we calculate Pearson correlation coefficient of equally weighted portfolio excess return with market excess return on large jackpot days and on non-large jackpot days, and the increase in correlation on large jackpots. Displayed are results from OLS regression of increase in correlation on large jackpots on the ranking of sorting values (1), and (2), respectively.

For the change in adjusted R-squared, we run the following regression for each portfolio on large jackpot days and non-large jackpot days:

$$PortRet_{j,t} = \alpha_j + \beta_j MktRet_t + \varepsilon_{i,t}$$

Where $PortRet_{i,t}$ is excess return of portfolio j at time t and $MktRet_t$ is market excess return at time t. Displayed are result from OLS regression of the increase of adjusted R-squared on large jackpots on the ranking of sorting values (1), and (2), respectively. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table A7
Individual firm's return co-movement with lottery-type and non-lottery-type portfolio return

		Co-movement with return of lottery-type stocks			Co-movement with return of non-lottery-type stocks		
		Large jackpot days	Non-large jackpot days	Difference (p value)	Large jackpot days	Non-large jackpot days	Difference (p value)
(1) Correlation coefficient	Mean	0.514	0.500	0.014*** (<.0001)	0.511	0.500	0.011*** (<.0001)
	Median	0.547	0.520	0.022*** (<.0001)	0.542	0.521	0.018*** (<.0001)
(2) Adjusted R-squared	Mean	0.282	0.261	0.021*** (<.0001)	0.278	0.260	0.018*** (<.0001)
	Median	0.296	0.271	0.021*** (<.0001)	0.291	0.271	0.017*** (<.0001)
Number of firms		817	817		817	817	

We firstly calculate idiosyncratic skewness in the past 150 trading days for each firm-date observation. Then we divide the ordered distribution of stocks by the skewness into three parts, each respectively representing the bottom 30%, 70% and the top 30% of the sample and calculate the equally weighted portfolio excess returns. The top 30% of the sample are considered as lottery-type stocks and the bottom 30% of the sample are considered as non-lottery-type stocks.

(1) Correlation coefficient: we calculate time-series Pearson correlation of individual stock return and lottery-type/non-lottery-type portfolio return (all adjusted by daily risk free rate, which is calculated from rate on central bank one-month time deposit) for large jackpot days and non-large jackpot days separately. For each firm, we obtain the paired difference in correlation calculated from large and non-large jackpot days. Reported are mean and median of correlation coefficient, and mean and median of the change on large jackpot days compared with non-large jackpot days among 817 firms.

(2) Adjusted R-squared: for large jackpot days and non-large jackpot days separately, we run the following regression for firm i to obtain adjusted R-squared:

$$Ret_{i,t} = \alpha_i + \beta_i PortRet_{i,t} + \varepsilon_{i,t}$$

Where $Ret_{i,t}$ is excess return of firm i at time t; $PortRet_{i,t}$ stands for portfolio excess return (lottery-type or non-lottery-type) at time t. Then we compute the paired difference of adjusted R-squared between large jackpot days and non-large jackpot days for each firm. Reported are mean and median of adjusted R-squared, and mean and median of the change on large jackpot days compared with non-large jackpot days among 817 firms.

Paired t test is employed for testing mean difference and Wilcoxon signed rank test is used for testing median difference. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

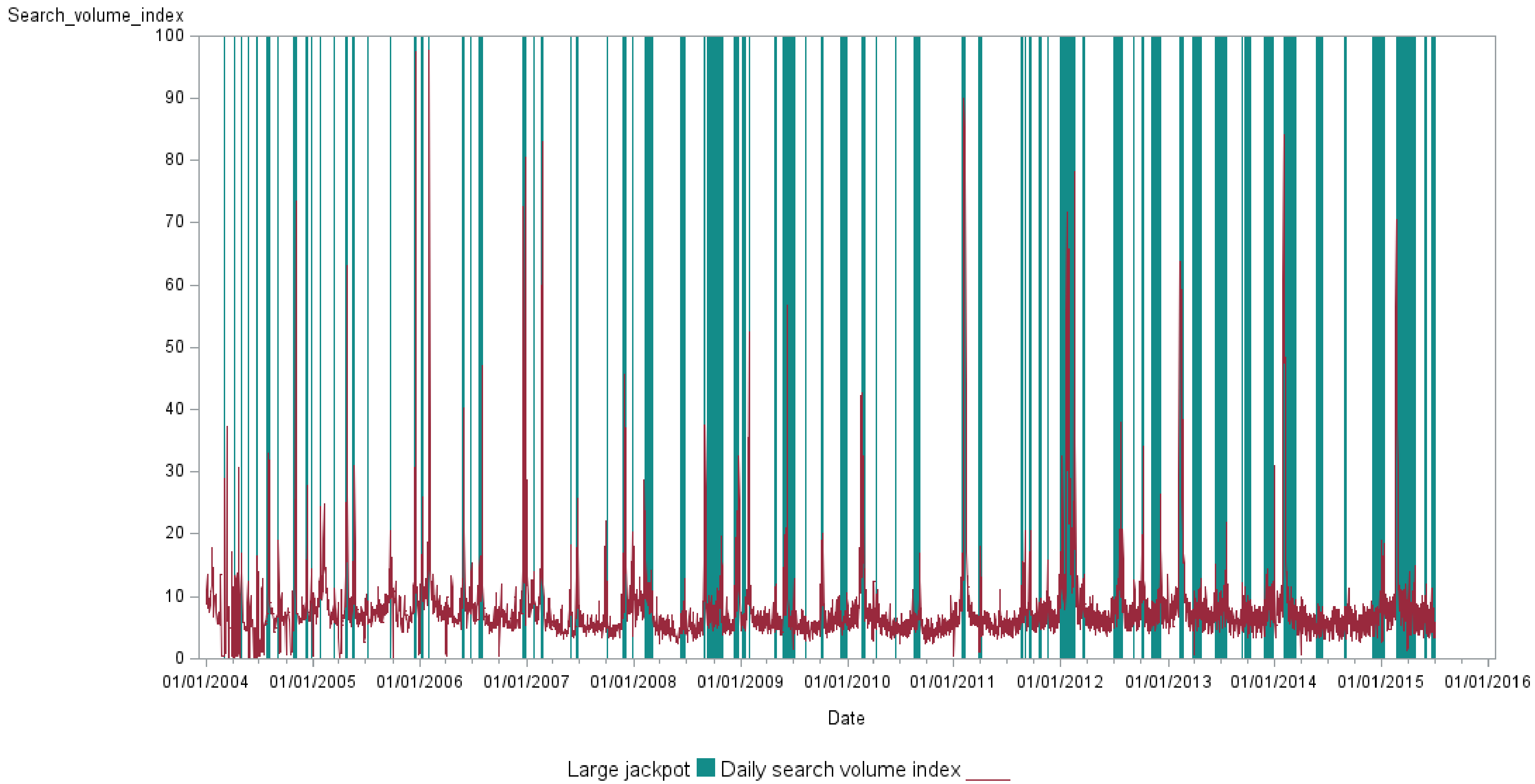


Figure A1. **Search volume index on large jackpot days and non-large jackpot days**

Plotted is the daily search volume index (red line) versus the dates (on the x-axis) and large jackpot days are highlighted in green. This figure shows that the Google daily search volume index experiences an increase on each large jackpot day. The sample period is from January 4, 2004 to June 30, 2015. We manually collect Google weekly search index over the sample period and unadjusted Google daily search index over each quarter. The search word is "樂透" which means "lottery" or "lotto" in English and search region is Taiwan. Since unadjusted daily search index is the relative search volume over each quarter, we calculate adjusted daily search index as follows: Adjusted daily search index= (Weekly search index to which the day belongs)*(Unadjusted daily search index/ Weekly average of unadjusted daily search index). After adjustments, the daily search volume index is bigger than 100 on 17 days. For clearer presentation of results, we adjust them back to 100 in this figure.