Almost won it all: The near-miss effect and future gambling activity

by

Jan Hanousek, Jr.
University of Memphis,
and Mendel University in Brno
Email: jan.hanousek12@gmail.com

Jan Hanousek
CERGE-EI, Charles University and the Academy of Sciences, Prague,
Mendel University in Brno, and CEPR London
Email: jan.hanousek@cerge-ei.cz

Jakub Mikulka
CERGE-EI, Charles University and the Academy of Sciences, Prague,
Email: jakub.mikulka@cerge-ei.cz

Jared Williams
University of South Florida
Email: jwilliams25@usf.edu

January 14, 2023

1. Introduction

In recent years, gambling-like activity has had a significant impact on the stock market. In the early 2020's, retail traders began actively trading "meme" stocks such as GameStop and AMC. These stocks' valuations rose far above what any realistic discounted cash flow analysis would justify, revealing significant limits to arbitrageurs' ability to keep stock market prices in line with fundamentals.[1] Because there was basically no relation between these stocks' price movements and news about their cash flows or market-wide discount rates, investing in these stocks was essentially a form of gambling. While it is widely accepted that the significant price run-ups were due to meme trading that had little relation to the firms' underlying cash flows, open questions remain regarding what causes retail traders to continue or quit their gambling activity, and thus, how long the meme stock mania is likely to persist. To rigorously examine these questions, a researcher would need a panel of investor/gambler trading activity that is large in the cross-section and long in the time series. In this paper, we analyze such a dataset and document the importance of the so-called "near-miss effect" on gamblers' propensity to continue gambling.

A "near miss" (also known as a "near hit" or "near win") is said to occur when a gambler almost, but did not, achieve a winning outcome. A near miss is not characterized by the magnitude of a player's loss, but rather, the realized state of the world being close (in some sense) to a state of the world in which a winning gamble would have been realized. For example, a "cherry–cherry–lemon" outcome on a slot machine can be considered a near miss, because "cherry-cherry-cherry" would have resulted in a large winning payout for the gambler. Regarding near misses and meme

[1] In April 2020, GameStop was trading at under $3 per share. At the end of 2020, it was trading at just under $20 per share. On January 27, 2021, GameStop closed at $347.51 per share, more than a 1,600% increase from its closing price on January 11, 2021. This dramatic price run-up was accompanied by abnormally large trading volume: From January 13-29, roughly 100 million GameStop shares traded each day, an increase of more than 14 times its 2020 average. For an SEC report on this, see: https://www.sec.gov/files/staff-report-equity-options-market-struction-conditions-early-2021.pdf

stock investing, a near miss could occur if a trader bought the stock too late, liquidated at just the wrong time (too early or too late), or if she was contemplating between two meme stocks and chose the wrong one. In terms of traditional fundamental value investing, there are many inputs that determine the future returns of a stock: the company's ability to innovate, the demand for its new products, the firm's ability to maintain its market share while its competitors are innovating, and regulatory risks. A near miss could occur if an investor were correct about all but one of these unknowns, but incorrect about one of them, causing a loss for the investor. Although it is widely believed that near misses cause gamblers to increase their gambling activity, prior research has arrived at different conclusions, presumably because the studies generally rely on small laboratory samples.[2]

In contrast to prior research on the near-miss effect, which typically relies on small laboratory samples, we analyze a unique, extensive panel data on both sports betting and lotteries provided by one of the largest betting companies in the Czech Republic. The dataset consists of more than 45 million sports and 40 million lottery tickets, covering more than 100,000 customers over seven years. The size and richness of the data allow us to analyze the long-term effects of a near miss on both the persistence of betting and the risk demanded by bettors. As Thaler and Ziemba (1988) note, sports betting markets have a very similar structure to conventional financial markets, making them suitable for analysis. Moreover, discount brokers and gambling sites use similar strategies to attract customers. McCabe (2021) shows that Robinhood started targeting college campuses with promotions and views it as a target demographic. This approach is consistent with the phenomenon described in gambling literature, which shows that the target

[2] Barton et al (2017) summarize the research on the near-miss effect as follows: "near misses were found to be associated with increasing one's bet, decreasing one's bet, or having no effect, each in a different study, making it difficult to determine whether near misses are capable of influencing per-play betting behaviour."

audience is single younger men (Barber and Odean, 2002; Andrikogiannopoulou and Papakonstantinou, 2020). A notable advantage of our setting is that betting odds are specified, which makes it is easier to determine people's risk tolerance and risk-seeking activity.

We find that both winning and a near miss are predictors of continued betting activity. While the house money effect can explain the relation between winning and future gambling activity, the relation between a near miss and future betting is unrelated to it. In sports betting, a near miss has a similar effect to winning along multiple dimensions: they both reduce future odds and the number of bets. Thus, as long as gamblers in the stock market can convince themselves that their losses are near misses, our findings suggest they will continue gambling on meme stocks even after suffering losses. However, near misses are distinct from winning in that they lead to a decrease (rather than increase) in the average amount of the bettor's future individual bets. This is consistent with the idea that individuals treat a near miss as a signal of future success, but they do not increase the average size of their bets because they do not have past winnings to spend. In contrast to sports betting, a near miss in lotteries is associated with a significant increase in the bettor's future average stake. In fact, the effect is even more impactful than winning large amounts, which suggests that individuals treat a near miss for lotteries as an incorrect signal of future success, which then leads to compulsive betting. Overall, the results show that a near miss, while still categorized as a loss, can lead to a very different reaction and impact future behavior differently.

Our paper is organized as follows: Section 2 summarizes previous research on the effects of past outcomes on future behavior and the near-miss effect and outlines the hypotheses of this paper. Section 3 describes the data and the sample creation. Section 4 describes the summary statistics and the results of our analysis. Finally, Section 5 concludes.

2. Literature review and hypotheses development

One of the reactions an individual has to a near-miss is increased persistence in playing (Cote et al., 2003). Interestingly, near-misses occur at a higher frequency than expected for commercial slot machines (Reid,1986; Harrigan, 2009), suggesting that commercial gambling system owners try to take advantage of the higher persistence in playing. Many previous studies document (e.g., Dixon and Schreiber, 2004) that cognitive reaction to near-miss differs from other types of losses, although the outcome is equivalent (i.e., both result in losing the initial bet). However, little is known about the near-miss effect observed in sports betting, which does not purely rely on luck. In a skill event, a close miss might provide helpful feedback to the individual and not be associated with negative emotion, as found by Sharman and Clark (2016). The main reason for the lack of research is that sports betting does not lead to near-miss by default, only when using particular types of bets called "*Accumulators*." Accumulator bets are a specific type of ticket that can have multiple wagers placed on it and will only pay off if all wagers win. It is important to note that there are no limits on the Accumulators, and many different wagers can be placed on them covering multiple sports and events. As a result, this allows us to specify and study the near-miss effect on behavior in the first hypothesis:

Hypothesis H1: Reaction to near miss loss differs from reaction to regular loss

Personality and behavior play an important role in decision-making for individuals, as documented by several studies (e.g., Cronqvist et al.,2012; Malmendier and Tate, 2008; Ham et al., 2017). Cronqvist et al. (2012) show that the CEO's personal leverage is a predictor of the leverage of a firm that (s)he manages. They refer to behavioral consistency theory, and their result supports the idea that individuals exhibit consistent behaviors across different situations and, when

given the opportunity, impose their preferences. Many studies support the concept of behavioral consistency (e.g., Epstein 1979 and 1980) and the idea that an individual's personality is stable and unchanging in adulthood (see, e.g., Costa & McCrae, 1988).

Moreover, individuals can be significantly affected by previous betting or investing outcomes, which has been documented in the literature (e.g., Barberis and Huang, 2001; Frydman et al., 2017; Kahneman and Tversky, 1979). Therefore, for correct identification, it may be necessary to study only the individuals for which the initial bet (investment) is observed. Otherwise, you may start observing them after a significant loss or a win, which may significantly impact their decision-making. Despite the rich literature on the house money effect, mental accounting, and prospect theory, most empirical literature ignores it and does not filter out individuals who started before the sample[3]. Missing information on previous individual wins or losses, or their omission, may skew the results.

Even recent studies of the betting market omit the fact that some of their users may have been gambling for a long time (e.g., Andrikogiannopoulou and Papakonstantinou, 2020). Clearly, the bettor's behavior and performance might be affected by previous wins or losses. In our paper's context, identifying the initial bet and using the subsample of bettors we have observed since the beginning is crucial. An individual placing a bet for the first time will consider his innate risk preference and will make a bet consistent with it. It will not be affected by his previous betting history. As a result, his first bet and its properties could proxy for his natural choice regarding risk. This leads us to specify the second hypothesis:

Hypothesis H2: Individuals' first bet proxies for the inherent preference of the user.

[3] See Thaler et al., 1990, Barberis and Huang, 2001, Frydman et al., 2017, among others.

One of the main concerns in past research was the data limitations. Many studies either use different data sources, which mimic individual choices for risky assets (e.g., mutual fund choices analyzed by Bailey et al., 2011), or employ laboratory experiments (e.g., Kahneman and Tversky, 1979). However, using any experimental design and data makes it difficult to answer research questions regarding behavioral patterns and long-term effects (e.g., Mental accounting, Disposition effect, etc.). The difficulties arise because of the short time horizon and the end-of-the-game problem. Usually, every experiment has only a limited number of rounds and explicitly or implicitly defines the end of the experiment (e.g., Frydman et al., 2018; Thaler and Johnson. 1990). Therefore, participants in these studies could behave differently than they would while participating in real betting or stock markets.

It is essential to mention that most past research does not account for individual behavioral traits and only studies the observed effects. For example, Thaler (1980, 1999) argues that individuals rebalance their portfolios based on their recent losses or gains. However, one of the implied assumptions is that the persistence of the different effects overpowers the behavior of the individual, and thus the investor makes portfolio rebalancing decisions only based on their past wins or losses. This assumption is made even though psychological research shows that behavior and personality traits, such as the Big Five personality traits[4], are very stable and even heritable (Costa and McCrae, 1994; Costa & McCrae, 1988). For example, Soldz and Vaillant (1999) conducted a 45-year study and found that personality traits are almost unchanged for the entire study duration and are considered plastic after the age of 30, despite a wide variety of shocks over the lifetime of the participants. It does not, therefore, seem reasonable to assume that there is not going to be any reversion back to the investment decisions based on personality, given the

[4] Openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism.

persistence of personality and "behavioral consistency," which shows that individuals follow the same principles throughout life (Funder and Colvin, 1991). Therefore, it is still unclear whether and how long the effects persist and how they would interact with the medium and long-term behavior of the investors. This leads us to specify the third and final hypothesis:

Hypothesis H3: Wins and losses will affect individuals only in a short-term horizon.

3. Data

We use a unique lottery and sports betting dataset, which has been provided to us by one of the largest betting companies in the Czech Republic. Our dataset spans January 2005 to February 2012, which consists of over 45 million sports tickets and over 40 million lottery tickets. Each ticket and bet have the following entries: time of the bet and time when the bet was resolved, stake, odds, winnings, and indicator of which betting opportunities on the ticket were winning or losing. Furthermore, we have an identifier of the customer and their gender, age, and location where the bet was placed[5], which can help us identify the municipality where the individual lives.

It is important to note that several users place many bets, primarily over the internet. The most likely motivation is to split larger bets into several smaller ones to avoid paying higher fees or keep lower odds than they would otherwise have. This element might be an important indicator since more professional bettors will likely only exhibit such behavior. As a result, we run our models on various restricted samples based on the maximum number of daily bets the individual placed during his entire betting history.

[5] While the bet could have been placed online or over the phone, where we do not have location information, vast majority of bets have been placed in person.

One of the most important aspects of betting is the possible combination of several bets on one ticket. While most tickets are *single,* (i.e., betting on just one event), more complicated types exist. Individuals could collect single bets as a list into so-called *combinator* ticket[6]. Trendy among the bettors were piling single bets into *accumulator (A)* tickets, where individual bets on several events on the same ticket and only wins if he/she correctly guesses all bets on the same ticket. The probability of winning the ticket is then the product of the implied probabilities associated with every item on the ticket. For accumulator tickets, we can then define the near miss when all but one bet on the ticket were winning. For completeness, there also exists a list of accumulator bets called *roster (R)* tickets, which we can treat as the list of independent accumulator tickets, because we know each accumulator's characteristics (odds, length, and near missed).

Accumulator bets offer much higher odds, but the expected winnings are lower than if the individual placed every bet individually. Furthermore, one could argue that accumulator bets are similar to option contracts since, to profit from option contracts, the investor essentially makes a simultaneous bet on both strike price and maturity.

Thanks to the richness of our dataset, we only consider individuals with clearly defined inception. Our sample frame started on January 1, 2005, and we excluded from the primary sample all individuals betting before April 1, 2005. We eliminate the potential influence of previous betting outcomes on an individual's decision-making by filtering out potentially active individuals before our data collection begins.

Lastly, it is necessary to note that the evaluation of intraday reactions is not straightforward. Specifically, the order of bets in the intraday segment might be close to random, especially those

[6] Because the total (ticket) odds for the combinator are hard to evaluate, there were not provided by the betting company. The success and relationship to the particular ticket odds could be analyzed using the stake and winning amount in case of win.

that differ by a few seconds or minutes. For example, if an individual places bets at a branch, the tickets could be placed in the system in a different or reverse order. As a result, capturing and evaluating immediate intra-day reactions is difficult. We, therefore, collapse the data to daily frequency and analyze the response from one day to the next. The summary statistics of our sample are available in Section 4.1.

4. Results

This section is organized as follows. First, in Section 4.1, we present summary statistics of the sample, outlining the decomposition and preference of users. Section 4.2 analyzes individuals' first three betting days to capture the determinants of continuation in betting activity. Lastly, Section 4.3 shows the dynamic effects on the sizes and frequency of betting.

Section 4.1

First, we outline the survival rate of new bettors in our sample. We report the distribution of participation (in days clusters) across age groups in Table 1.

---------- Insert Table 1 around here ----------

Table 1 above shows the distribution of sports betting participation across the age groups. We can see that the largest number of bettors is for the age group 20-30, with the number of individuals slowly decreasing as age increases. Many users tend to place only a couple of bets and leave the sample. Therefore, it may be important to understand what push and pull factors are for individuals to continue playing. We report the results of this analysis in Section 4.2. We can see that the most populated age category of longer-playing gamblers is between 30 to 40 years. Overall, we have over 70,000 new users, with more than 15,000 bettors staying in the sample for more than 31 days.

Let us also analyze the survival rate of individuals for lotteries, which we report in Table 2 below.

---------- Insert Table 2 around here ----------

We can see that the number of individuals who try the lottery is less than four times lower, and the exit from the lottery sample is much higher than for sports betting. We can also see that the overall survivor rate is lower for lotteries than for sports betting. It is in line with the general wisdom that lottery betting might be less active than sports betting, where individuals place bets while watching sporting events. Compared to that, people might only place lottery bets on specific "lucky" days and require lower participation. Interestingly, we can also see that the most populated category is 30-40. However, the percentage of users staying in the sample for longer is positively related to age and does not peak like sports betting.[7] Overall, the results confirm that the motivations for lotteries and sports betting might be significantly different.

However, it is also important to account for the risk sought by individuals, both in sports betting and lotteries. The results of the summary statistics are available in Table 3.

---------- Insert Table 3 around here ----------

First, we can see that for sports betting, the mean odds (stake) increase until they peak for individuals in the 40-50 (30-40) category, then decrease. This finding is consistent with Gervais and Odean (2001), who document that individuals tend to be less overconfident with increasing age. We observe a similar pattern for lotteries, with mean stake peaking for individuals between 40-50 and mean odds peaking for individuals in the 50-60 category. Observed patterns are in line

[7] Increased duration time with age in lottery betting could be in line with additional social interactions for older bettors. Lottery betting in our sample has been primarily restricted to betting venues (towns), and therefore, betting lottery tickets means going to the town.

with our expectations. Lottery outcomes are random and do not require particular skills; therefore, they could be more attractive to less confident or less sophisticated bettors. On the other hand, lottery-type rewards could be substantial and, therefore, more attractive to low-income and older people.

One important aspect of the data is the usage of accumulator bets. These bets are the riskiest type of bet since the likelihood of winning most accumulator bets is on par with a lottery ticket. As a result, it is important to analyze what is the determinant of individuals using accumulator bets. We can see that the accumulator bets are far more popular for sports compared to lotteries and are roughly twice as popular. Another interesting observation is that accumulator bets are far more popular for younger individuals, with 80% of observations for people below 30 years of age being for accumulator bets. The average stake for accumulator sports bets is also much larger despite the much lower chance of winning. This suggests that for many individuals, an accumulator sports bet might be similar to a lottery in payout, but unlike the lottery, its outcome is not purely random based. This observation may be consistent with the results of Kumar (2009), who shows that individual investors prefer stocks with lottery-like features. Conversely, this does not seem to be the case for lotteries, where only 40% of bets are part of accumulators, and the mean stake for accumulator bets is less than half of the non-accumulator tickets.

These results suggest that accumulator sports bets provide an essential feature for individual gamblers since they are not purely random, like lotteries, but offer lottery-like returns. Since our sample contains the 2008 crisis, the increased preference might be the result of the increased demand for lottery-like returns, as shown by Kumar (2009).

To conclude this analysis, we need to analyze the decision never to use accumulator bets. We provide the results of the summary table in Table 4.

---------- Insert Table 4 around here ----------

We can see that more than 90% of all users tried accumulator bets at least once for sports gambling. Furthermore, we see a monotone pattern in the number of users never trying accumulator bets, where increasing age decreases the likelihood of ever trying accumulator bets. Conversely, the number of users never trying accumulator bets in lotteries is above 20%. Other than the youngest category, the percentage of individuals who never try accumulator bets in lotteries is comparable across age groups, suggesting that the type of individual to try accumulator bet in lotteries seems to be constant.

Section 4.2

As outlined in Section 3, capturing the initial bet of individuals is necessary since the result of the first bet can influence the continuation of betting activity. Furthermore, analysis of the initial bet allows us to observe individual reactions that are not affected by previous losses or winnings. This should enable us to capture an unbiased response. We, therefore, analyze the decision to continue after each of the first three days. The analysis of whether the individual will continue after the first day is available in Table 5.

---------- Insert Table 5 around here ----------

The dependent variable in Table 5 is a dummy variable of whether the individual continues after the first day of betting. We analyze sports betting and lotteries, and for every model we report the coefficients for the entire sample and a subsample for users who placed less than 200 bets on the day. We can see that a net win[8] on day one increases the likelihood that an individual will continue

[8] Winnings are larger than the stake plus the fee to place the bet (fee for a small and regular bets for venue and internet betting were 5% and 10%, respectively).

to gamble. This finding is consistent with the house money hypothesis since individuals might treat the money they won differently and use it as capital in future bets. Interestingly, suffering near misses also increases the likelihood of the user continuing to gamble in the future. However, this is not observed for lotteries where near miss is not a significant predictor. This result is important, since it suggests that previous studies focusing on winnings and losses might have missed this critical element. Even though the individual lost, if he lost by a near miss or something he might perceive as a near miss, the response might be more akin to winning than losing.

This result can be surprising. In a systematic review, Barton et al. (2017) show that near miss motivates game continuation. However, it is essential to note the distinction between near misses in sports gambling, which can be perceived as skill-based activity, compared to random lotteries. In a skill-based event, a near miss might give helpful feedback and encourage the player that success is within reach (Reid,1986). This might not be the case in lotteries, which our results suggest.

Another interesting result is when players lose big amounts[9] in the initial bet. We observe that this generally leads to sports bettors stopping playing. However, it has the opposite effect for lottery players, who tend to be much more likely to stay gambling for the next period, perhaps hoping to win it back. This behavior is consistent with prospect theory and disposition effect.

Lastly, an important observation is that many individuals start their betting career using the riskiest asset, i.e., the *accumulator* ticket. As seen in Tables 4 and 5, many users start with the accumulator ticket; for the vast majority, it is the primary type of bet they use. It also seems to be affected by age since we observe a drop in the frequency of accumulator tickets with increasing age. We show in this table that if individuals start with accumulator tickets in sports betting

[9] We define a large amount as more than 2,000 CZK, about 1/10 the average monthly wage in the Czech Republic at that time.

(lotteries), it increases (decreases) their likelihood of continuing betting. This result may be surprising since it suggests that hazardous individuals might have high persistence in gambling, and they might immediately start with the riskiest asset.

Let us continue the analysis by analyzing what are the determinants of users continuing after the second day of gambling. The results of this analysis are available in Table 6.

---------- Insert Table 6 around here ----------

The dependent variable in Table 6 is a dummy variable of whether the individual continues after the second day of betting. Again, we report results for both sports and lotteries and report estimations for both the entire sample and a sample with users that placed less than 200 bets on a given day. First, we again see that net win, on both the first and second day, is a significant and positive predictor of future gambling. This is again consistent with the house money hypothesis; however, it is important to note that the net win on the second day has a much larger effect than winning on day one. This result is related to our hypothesis H3 since it suggests that recent events will impact the individual more. We see this effect repeated for both near misses and significant losses. We also see that a near miss on the second day is a significant positive predictor for the continuation of gambling for sports betting. Furthermore, the estimated effect is comparable to the previous Table 5. This result again stresses the importance of accounting for the near-miss effect. While a near miss is still categorized as a loss, we can see its impact comparable to winning on the previous day. This again shows that previous studies, which study the effects of wins and losses, might have omitted a substantial factor.

Interestingly, we observe that a near miss on the second day leads to a lower chance of continuing in lotteries. This result goes against the prevailing sentiment that near miss motivates future play (e.g., Barton et al., 2017). It is important to note that there are far fewer individuals

betting on lottery accumulators, meaning that the estimate might be influenced by the fewer users using accumulators for lotteries. Furthermore, unlike lottery machines, the near miss might be more difficult to capture for lotteries. As Kahneman and Tversky (1982) pointed out for lotteries, the near miss with the winning number might depend on the ease with which the winning outcome can be imagined[10]. This idea is supported by a study by Sundali et al. (2012), who do not find that near miss leads to increased playing for roulette players.

Lastly, an important effect for lotteries seems to be a significant loss on the second day, which motivates future play. This is again consistent with prospect theory since players are encouraged to make money back. We do not observe this effect for sports betting, nor does winning significantly impact the future play for both lotteries and sports betting.

To finish our analysis of the beginning of a gambler's career, we analyze the determinants of continuing on the third day of gambling. This analysis is available in Table 7.

---------- Insert Table 7 around here ----------

The dependent variable in Table 7 is a dummy variable of whether the individual continues after the third day of betting. Table 7 follows the format of Tables 5 and 6. We again find that net win motivates the user to continue playing for both lotteries and sports gambling, consistent with the house money effect. Furthermore, consistent with our previous result and hypothesis, we again see that the impact of net win is most significant for the most recent day, with the last day's effect decreasing in magnitude.

We again observe that near miss significantly affects continuing to gamble for sports betting. Similar to net winning, the effect is most substantial for the near miss on day 3, with the

[10] Kahneman and Tversky (1982) illustrate on a case of lottery with 865304 as the winning number with three clients holding the ticket numbers of 361204, 965304 and 865305. They note that these numbers are ordered in the degree of disappointment, with the last number being the most impactful in terms of near miss effect.

magnitude dropping for previous days. This is consistent with our hypotheses H1 and H3 and stresses the importance of accounting for near misses in future studies. Interestingly, we do not observe any near-miss effect for lotteries. This finding suggests that similar to roulette gambling, the near-miss effect in lotteries might be more difficult to capture (Sundali et al., 2012).

Lastly, we again observe that significant wins do not seem to impact future gambling for both lotteries and sports gambling. Similarly, substantial losses do not show a discernible pattern for sports betting and are not significant for lotteries.

Section 4.3

While the previous section analyzes the first three gambling days and the determinants of play continuation, it does not analyze what affects the risk demand of players. This section analyzes the effects of winning, losing, and a near miss on the demanded odds, the stake, and the number of bets. To avoid overlap with our previous analysis, we analyze the betting days starting from day four. Furthermore, to prevent the end-of-the-game problems, i.e., the possibility that a user changes his betting strategy when he nears his exit from the game, we end the analysis four days before his last day as a gambler in our sample. The result of this analysis for sports betting is available in Table 8.

---------- Insert Table 8 around here ----------

The dependent variables are the mean odds, mean stake, and the number of bets recorded as the natural logarithm for the given day. There are two sets of dependent variables. First, as outlined in the previous Section 4.2, it is important to account for the initial bet, which is why we use the key variables from the first day as predictors of all future behaviors of the individual. Second, we use the same variables but constructed from the previous betting day.

We can see that net winning motivates future play at the beginning of the career. It adversely impacts future odds and the number of bets while increasing the mean stake. This suggests that the individual starts preferring less risky bets while increasing the amount they bet. Interestingly, the net win on the first day is significant and in the same direction as winning on the previous day. However, the effects of the last day dominate the effects from day one. This result is somewhat consistent with the house money hypothesis since it shows that winning increases gambling activity as the average stake betted increases. However, it is also consistent with prospect theory (e.g., Kahneman and Tversky, 1979), since the user exhibits risk aversion in terms of the odds and the number of bets.

Another important effect is again the near miss. We can see that a near miss on the previous day reduces the mean odds, similar to winning, but decreases the stake and increases the number of bets instead of winning. This result again shows that despite being classified as a loss, near miss leads to different reactions compared to winning or losing. These results suggest that the individual will treat a near miss as almost having won, which leads to reducing the mean odds, but he reduces the stake since he has no winnings that he could use as capital. The increase in the number of bets can be seen as compulsive gambling and a potential desire to break even, which is again consistent with prospect theory. Interestingly, the near miss from the first day has a persisting effect, although lower in magnitude compared to the current near miss.

Lastly, we can observe the effects of large wins and losses. Compared to a net win, we can see that a large win has similar and intensifying effects. However, the response differs between the full sample, including large professional bettors, and the sample of only regular users. Regular users seem to respond to large wins similarly to regular wins, with the biggest difference being a much larger increase in the mean stake, which is consistent with the house money hypothesis.

Compared to that, professional users seem to respond by also increasing the mean odds and number of bets following a large win. This may be the result of intensifying the house money effect and more compulsive gambling. Interestingly, a large loss seems to have a similar effect to a large win, with regular bettors also decreasing the mean odds and number of bets but increasing the stake. However, the response for large loss seems to be higher in terms of magnitude. This may be a result of the desire to break even, which is again consistent with prospect theory and the disposition effect.

It is again important to analyze the reaction of lottery players since analysis in Section 4.2 showed that it may significantly differ from sports gambling. This analysis is available in Table 9.

---------- Insert Table 9 around here ----------

We can see that similar to sports betting, winning leads to a decrease in mean odds and an increase in the mean stake. However, the number of bets also increases for lottery players. This result shows that for lottery players, the house money effect might dominate since the increase in the number of tickets is consistent with more impulsive gambling.

A very important difference, however, is the reaction to a near miss. We can see that while regular lottery players do not seem to change their mean odds, they do seem to increase the mean stake and the number of bets. This is consistent with near miss literature (CITE), since it suggests that near miss in random-based gambling leads to more compulsive betting. Interestingly, when taken together with the results from the previous section, it suggests that the response to near miss differs based on when the player suffers it in their lifetime. Suffering a near miss in random-based gambling, such as lotteries, decreases the chance of continuing playing in the beginning. However, once a player is more established, a near miss tends to lead to more compulsive gambling. Very important is also the distinction between a near miss in sports betting, which is more skill-based

and more akin to financial markets (Thaler and Ziemba,1988), and lotteries. For sports betting, near miss has a similar effect to winning but differs in that the mean stake is lower since the user does not have the winnings to bet. For lotteries, the prospect theory and the desire to break even seem to prevail. The near-miss leads to an increase in the stake, which is the opposite of sports gambling. This might suggest that lottery players might incorrectly treat a near miss as an indicator of their skill and a false signal that they are close to winning. This might be one of the reasons why a near miss are programmed to occur more frequently on slot machines (Reid,1986; Harrigan, 2009).

Lastly, the lottery results suggest that the house money effect might be stronger than sports gambling. The lottery players also increase the mean odds following significant wins, compared to sports gambling.

5. Conclusion

Motivated by the recent wave of meme stock trading in the stock market, we analyze a unique, extensive panel dataset consisting of more than 45 million sports and 40 million lottery tickets, covering more than 100,000 customers over seven years, to examine the effects of a near miss on future gambling activity. We find that near misses, like winning bets, increase the likelihood that a bettor will continue to gamble. Thus, as long as gamblers in the stock market can convince themselves that their losses are near misses, our findings suggest they will continue gambling on meme stocks even after suffering losses. However, when we look at other dimensions of gambling such as a gambler's average future betting stake, the effects of a near miss differ from those of a winning bet. Exploring the distinctions between near misses and wins/losses using financial data is an interesting avenue for future research.

References

Andrikogiannopoulou, A., & Papakonstantinou, F. (2020). History-dependent risk preferences: Evidence from individual choices and implications for the disposition effect. *The Review of Financial Studies*, *33*(8), 3674-3718.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *The Journal of Finance*, *77*(6), 3141-3190.
Barber, B., and T. Odean (2002). Online investors: Do the slow die first? *Review of Financial Studies* 15:455–87.

Barberis, N., & Huang, M. (2001). Mental accounting, loss aversion, and individual stock returns.T*he Journal of Finance*, 56(4), 1247-1292.

Barton, K. R., Yazdani, Y., Ayer, N., Kalvapalle, S., Brown, S., Stapleton, J., ... & Harrigan, K. A. (2017). The effect of losses disguised as wins and near misses in electronic gaming machines: A systematic review. *Journal of Gambling Studies*, *33*(4), 1241-1260.

Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), Can personality change? (pp. 21–40). American Psychological Association. https://doi.org/10.1037/10143-002

Côté, D., Caron, A., Aubert, J., Desrochers, V., & Ladouceur, R. (2003). Near wins prolong gambling on a video lottery terminal. *Journal of Gambling Studies*, 19(4), 433–438.

Cronqvist, H., Makhija, A. K., & Yonker, S. E. (2012). Behavioral consistency in corporate finance: CEO personal and corporate leverage. *Journal of Financial Economics*, 103(1), 20-40.

Dixon, M. R., & Schreiber, J. E. (2004). Near-miss effects on response latencies and win estimations of slot machine players. *Psychological Record*, 54(3).

Epstein, S., 1979. The stability of behavior: I. on predicting most of the people much of the time. *Journal of Personality and Social Psychology 37*, 1097-1126.

Epstein, S., 1980. The stability of behavior. *American Psychologist* 35, 790-806.

Frazzini, A. (2006). The disposition effect and underreaction to news. The Journal of Finance, 61(4), 2017-2046.

Frydman, C., Hartzmark, S. M., & Solomon, D. H. (2018). Rolling mental accounts. *The Review of Financial Studies*, 31(1), 362-397.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *Review of Financial Studies*, 14(1), 1-27.

Ham, C., Seybert, N., & Wang, S. (2018). Narcissism is a bad sign: CEO signature size, investment, and performance. *Review of Accounting Studies*, *23*(1), 234-264.

Harrigan, K. A. (2009). Slot machines: Pursuing responsible gaming practices for virtual reels and near misses. *International Journal of Mental Health and Addiction*, 7(1), 68–83.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47, 263–91.

Kahneman, D. & Tversky, A. (1982). The Psychology of Preferences. *Scientific American, January*, 136-142.

Kumar, A. (2009). Who gambles in the stock market?. *The Journal of Finance*, 64(4), 1889-1933.
Malmendier, U., & Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, *89*(1), 20-43.

McCabe, C., 2021. It isn't just AMC. Retail Traders Increase Pull on the Stock Market Retrieved from https://www.wsj.com/articles/it-isnt-just-amc-retail-traders-increase-pull-on-the-stock-market-11624008602.

Ozik, G., Sadka, R., & Shen, S. (2021). Flattening the illiquidity curve: Retail trading during the COVID-19 lockdown. *Journal of Financial and Quantitative Analysis*, *56*(7), 2356-2388.

Reid, R. L. (1986). The psychology of the near miss. Journal of Gambling Behavior, 2(1), 32–39.
Sharman, S., & Clark, L. (2016). Mixed emotions to near-miss outcomes: A psychophysiological study with facial electromyography. *Journal of Gambling Studies*, *32*(3), 823-834.

Sundali, J. A., Safford, A. H., & Croson, R. (2012). The impact of near-miss events on betting behavior: An examination of casino rapid roulette play. *Judgment and Decision making*, *7*(6), 768-778.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of economic behavior & organization*, *1*(1), 39-60.

Thaler, R. H., Ziemba, W. T., 1988. Anomalies: Parimutuel betting markets: Racetracks and lotteries. *The Journal of Economic Perspectives* 2 (2), 161–174.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, *36*(6), 643-660.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making*, *12*(3), 183-206.

Table 1. Distribution of participation of new bettors on SPORTS betting across the age groups.

The first number on each cell corresponds to the total number of bettors; the second is the row percentage. The table contains a distribution for user participation, where we exclude users with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies).

Age group	Number of bettors: Distribution by number of days across age groups						
	<4	4-10	11-20	21-30	31-100	>100	Total
<20	910	843	571	254	332	8	2,918
	31.19	28.89	19.57	8.7	11.38	0.27	100
20-30	7,165	5,948	4,413	2,407	4,952	386	25,271
	28.35	23.54	17.46	9.52	19.6	1.53	100
30-40	4,276	3,712	3,198	1,869	4,513	524	18,092
	23.63	20.52	17.68	10.33	24.94	2.9	100
40-50	2,967	2,436	2,040	1,234	3,141	448	12,266
	24.19	19.86	16.63	10.06	25.61	3.65	100
50-60	1,927	1,220	918	568	1,650	287	6,570
	29.33	18.57	13.97	8.65	25.11	4.37	100
>60	1,155	679	545	369	922	213	3,883
	29.75	17.49	14.04	9.5	23.74	5.49	100
missing	1,021	287	147	80	168	16	1,719
	59.39	16.7	8.55	4.65	9.77	0.93	100
Total	19,421	15,125	11,832	6,781	15,678	1,882	70,719
	27.46	21.39	16.73	9.59	22.17	2.66	100

Note: We excluded all established users and analyzed only those with the first bet appearing after April 1, 2005

Table 2. Distribution of participation of new bettors on LOTTERY betting across the age groups.

The first number on each cell corresponds to the total number of bettors; the second is the row percentage. The table contains a distribution for user participation, where we exclude users with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies).

Age group	Number of bettors: Distribution by number of days across age groups						
	<4	4-10	11-20	21-30	31-100	>100	Total
<20	149	42	18	2	7	0	218
	68.35	19.27	8.26	0.92	3.21	0	100
20-30	6,574	1,778	870	396	493	14	10,125
	64.93	17.56	8.59	3.91	4.87	0.14	100
30-40	6,757	2,194	1,301	574	889	49	11,764
	57.44	18.65	11.06	4.88	7.56	0.42	100
40-50	4,574	1,736	1,135	513	884	63	8,905
	51.36	19.49	12.75	5.76	9.93	0.71	100
50-60	2,523	1,094	725	385	713	35	5,475
	46.08	19.98	13.24	7.03	13.02	0.64	100
>60	1,709	676	471	249	524	31	3,660
	46.69	18.47	12.87	6.8	14.32	0.85	100
missing	1,035	356	171	82	122	10	1,776
	58.28	20.05	9.63	4.62	6.87	0.56	100
Total	23,321	7,876	4,691	2,201	3,632	202	41,923
	55.63	18.79	11.19	5.25	8.66	0.48	100

Note: We excluded all established users and analyzed only those with the first bet appearing after April 1, 2005

Table 3. Basic descriptive statistics: Mean and median distributions for betting characteristics across age groups.

Columns correspond to daily characteristics by individuals actively betting for more than ten days, excluding the potential use of bots. The first number on each cell stands for the total number of observations (days*bettor); the second is the mean, and the third is the median of the particular variable on the age group. Variable of interest are: mean daily odds (by bettor), mean stake, daily stake on Accumulator-type bets, and dummy equal to one for using an Accumulator bet on a given day.

	Sports				Lottery			
	Mean(odds)	Mean(stake)	Stake(AR)	AR dummy	Mean(odds)	Mean(stake)	Stake(AR)	AR dummy
<20	27,520	30,445	30,445	30,445	208	545	545	545
	162.5	59.8	88.2	0.8	8,833.2	25.3	12.1	0.4
	7.2	22.0	30.0	1.0	650.0	20.0	-	-
20-30	365,236	421,054	421,054	421,054	25,552	46,256	46,256	46,256
	380.4	143.3	162.5	0.8	10,703.0	40.5	15.7	0.5
	9.4	33.3	30.0	1.0	253.3	20.0	-	-
30-40	320,528	393,730	393,730	393,730	38,931	80,591	80,591	80,591
	617.7	254.3	272.5	0.7	12,493.2	46.5	13.9	0.4
	10.5	50.0	40.0	1.0	650.0	24.0	-	-
40-50	219,555	286,085	286,085	286,085	34,262	80,388	80,388	80,388
	813.3	201.5	207.6	0.7	12,318.7	49.3	12.6	0.4
	11.8	47.5	30.0	1.0	650.0	30.0	-	-
50-60	109,480	154,110	154,110	154,110	24,264	60,195	60,195	60,195
	691.8	121.6	111.5	0.7	14,439.7	36.4	10.8	0.4
	14.6	40.0	20.0	1.0	650.0	27.3	-	-
>60	62,887	95,191	95,191	95,191	18,882	43,531	43,531	43,531
	789.5	110.3	72.3	0.6	10,376.9	34.8	10.4	0.4
	17.7	30.0	20.0	1.0	352.5	20.0	-	-
missing	9,534	15,101	15,101	15,101	2,293	11,354	11,354	11,354
	492.3	120.2	108.9	0.6	18,847.2	39.7	5.0	0.2
	21.2	41.3	20.0	1.0	650.0	30.0	-	-
Total	1,114,740	1,395,716	1,395,716	1,395,716	144,392	322,860	322,860	322,860
	583.1	179.8	188.8	0.7	12,281.0	42.6	12.5	0.4
	10.8	40.0	30.0	1.0	565.0	24.0	-	-

Table 4. Distribution of bettors never using Accumulator-type bets across age groups

The first number on each cell corresponds to the total number of bettors; the second is the row percentage. The table contains distribution for users participating on the betting for more than ten days, where we exclude users with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies).

Age group	Using A-type bets on SPORTS			Using A-type bets on LOTTERY		
	Sometime	Never	Total	Sometime	Never	Total
<20	2,821	97	2,918	87	131	218
	96.68	3.32	100	39.91	60.09	100
20-30	24,382	889	25,271	7,954	2,171	10,125
	96.48	3.52	100	78.56	21.44	100
30-40	17,279	813	18,092	9,511	2,253	11,764
	95.51	4.49	100	80.85	19.15	100
40-50	11,680	586	12,266	6,938	1,967	8,905
	95.22	4.78	100	77.91	22.09	100
50-60	6,090	480	6,570	4,269	1,206	5,475
	92.69	7.31	100	77.97	22.03	100
>60	3,579	304	3,883	2,927	733	3,660
	92.17	7.83	100	79.97	20.03	100
missing	461	1,258	1,719	558	1,218	1,776
	26.82	73.18	100	31.42	68.58	100
Total	66,292	4,427	70,719	32,244	9,679	41,923
	93.74	6.26	100	76.91	23.09	100

Note: We excluded all established users and analyzed only those with the first bet appearing after April 1, 2005

Table 5. Continue betting after the first day

The dependent variable is a dummy variable equal to one if the individual continues betting after the first day. The unit of observation is the (new) user, starting after April 1, 2005. It has been estimated by the linear probability model (LPM); the estimated coefficients are, therefore, directly linked to the probabilities of continuing on the game. Coefficients are similar to marginal effects on the mean computed from a binary choice model calculated by the delta methods. Alternative estimation methods results and detailed coefficients for age and year fixed effects are available upon request or on the Internet Appendix. Columns marked as Full contaon the results for the entire sample of new users. On the other columns, we exclude bettors with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Heteroskedasticity consistent standard errors are on brackets.

Variables	Sports bets				Lotteries			
	Full	<200	Full	<200	Full	<200	Full	<200
=1 if the (net) win on day 1	0.051***	0.064***	0.052***	0.065***	0.119***	0.129***	0.115***	0.116***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.011)	(0.013)	(0.011)	(0.013)
=1 if near-missed on day 1	0.015***	0.015***	0.013***	0.013***	-0.162	-0.233	-0.154	-0.21
	(0.003)	(0.004)	(0.003)	(0.004)	(0.202)	(0.237)	(0.202)	(0.237)
=1 if loss>2,000CZK on day 1	-0.015*	-0.026**	-0.014	-0.025**	0.198***	0.210**	0.195***	0.198**
	(0.009)	(0.011)	(0.009)	(0.011)	(0.060)	(0.083)	(0.060)	(0.083)
=1 if netwin>2,000CZK on day 1	0.009	0.016	0.008	0.016	0.073	0.045	0.076*	0.054
	(0.009)	(0.012)	(0.009)	(0.012)	(0.046)	(0.065)	(0.046)	(0.065)
=1 if A or R bet on day 1			0.009***	0.011***			-0.015***	-0.034***
			(0.002)	(0.003)			(0.005)	(0.006)
=1 if a man, female is the base category	0.081***	0.076***	0.081***	0.076***	-0.050***	-0.031***	-0.051***	-0.032***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.007)	(0.006)	(0.007)
Age group FE, year FE	yes	yes	yes	yes	yes	yes	yes	yes
R^2, adjusted	0.047	0.042	0.047	0.042	0.042	0.029	0.042	0.03
N (Observations)	91,585	70,719	91,585	70,719	49,637	41,923	49,637	41,923

Table 6. Continue betting after the second day

The dependent variable is a dummy variable equal to one if the individual continues betting after the second day. The unit of observation is the (new) user, starting after April 1, 2005. It has been estimated by the linear probability model (LPM); the estimated coefficients are, therefore, directly linked to the probabilities of continuing on the game. Age group (10 years) fixed effect, and year FE were used as controls. Alternative estimation methods results and detailed coefficients for age and year fixed effects are available upon request or on the Internet Appendix. Columns marked as Full contain the results for the entire sample of new users. In the other columns, we exclude bettors with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Heteroskedasticity consistent standard errors are on brackets.

Variables	Sports bets				Lotteries			
	Full	<200	Full	<200	Full	<200	Full	<200
=1 if (net) win on day 1	0.013***	0.016***	0.012***	0.015***	0.023**	0.020*	0.021**	0.012
	(0.002)	(0.003)	(0.002)	(0.003)	(0.009)	(0.012)	(0.009)	(0.012)
=1 if (net) win on day 2	0.062***	0.080***	0.062***	0.079***	0.071***	0.088***	0.070***	0.083***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.008)	(0.011)	(0.008)	(0.011)
=1 if near-missed on day 1	0.002	-0.000	0.003	0.001	0.564**	0.512*	0.566**	0.521*
	(0.003)	(0.004)	(0.003)	(0.004)	(0.247)	(0.301)	(0.247)	(0.301)
=1 if near-missed on day 2	0.016***	0.018***	0.016***	0.018***	-0.624***	-0.631**	-0.621***	-0.620**
	(0.003)	(0.004)	(0.003)	(0.004)	(0.214)	(0.246)	(0.214)	(0.246)
=1 if loss>2,000CZK on day 1	0.002	-0.001	0.002	-0.001	-0.002	-0.033	-0.003	-0.038
	(0.009)	(0.012)	(0.009)	(0.012)	(0.052)	(0.081)	(0.052)	(0.081)
=1 if loss>2,000CZK on day 2	0.015*	0.017	0.014*	0.017	0.121***	0.142**	0.119***	0.135*
	(0.008)	(0.012)	(0.008)	(0.012)	(0.044)	(0.071)	(0.044)	(0.071)
=1 if netwin>2,000CZK on day 1	0.003	0.005	0.003	0.005	0.100***	0.139**	0.102***	0.144**
	(0.009)	(0.012)	(0.009)	(0.012)	(0.038)	(0.059)	(0.038)	(0.059)
=1 if netwin>2,000CZK on day 2	-0.010	-0.010	-0.010	-0.011	0.058	0.074	0.059*	0.077
	(0.009)	(0.012)	(0.009)	(0.012)	(0.035)	(0.054)	(0.035)	(0.054)
=1 if A or R bet on day 1			-0.005**	-0.007**			-0.006	-0.018***
			(0.002)	(0.003)			(0.005)	(0.006)
=1 if A or R bet on day 2			0.002	0.002			-0.002	-0.008
			(0.002)	(0.003)			(0.004)	(0.006)
=1 if a man, female is the base category	0.067***	0.067***	0.067***	0.067***	-0.025***	-0.012*	-0.026***	-0.013*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.007)	(0.005)	(0.007)
R^2, adjusted	0.041	0.036	0.041	0.036	0.025	0.017	0.025	0.017
N (Observations)	82790	61924	82790	61924	34342	26628	34342	26628

Table 7. Continue betting after the third day

The dependent variable is a dummy variable equal to 1 if the individual continues betting after the 3rd day. The sample and controls are the same as on Table 5-6.

	Sports bets				Lotteries			
	Full	<200	Full	<200	Full	<200	Full	<200
=1 if (net) win on day 1	0.006***	0.008***	0.006***	0.007**	0.015*	0.013	0.012	0.006
	(0.002)	(0.003)	(0.002)	(0.003)	(0.008)	(0.012)	(0.009)	(0.012)
=1 if (net) win on day 2	0.021***	0.027***	0.021***	0.027***	0.027***	0.034***	0.025***	0.029***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.007)	(0.010)	(0.008)	(0.010)
=1 if (net) win on day 3	0.058***	0.078***	0.058***	0.078***	0.056***	0.072***	0.055***	0.068***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.008)	(0.010)	(0.008)	(0.010)
=1 if near-missed on day 1 (by one success)	0.003	0.003	0.005*	0.005	0.134	0.155	0.139	0.166
	(0.003)	(0.004)	(0.003)	(0.004)	(0.300)	(0.343)	(0.300)	(0.343)
=1 if near-missed on day 2 (by one success)	0.008***	0.008**	0.008***	0.008**	-0.086	-0.013	-0.085	-0.016
	(0.003)	(0.003)	(0.003)	(0.004)	(0.376)	(0.485)	(0.376)	(0.485)
=1 if near-missed on day 3 (by one success)	0.017***	0.019***	0.017***	0.019***	0.107	0.162	0.103	0.161
	(0.003)	(0.004)	(0.003)	(0.004)	(0.161)	(0.200)	(0.161)	(0.200)
=1 if loss>2,000CZK on day 1	0.007	0.006	0.007	0.006	0.012	0.008	0.011	0.004
	(0.008)	(0.012)	(0.008)	(0.012)	(0.047)	(0.078)	(0.047)	(0.078)
=1 if loss>2,000CZK on day 2	0.030***	0.041***	0.029***	0.041***	0.062	0.081	0.060	0.075
	(0.008)	(0.012)	(0.008)	(0.012)	(0.040)	(0.068)	(0.040)	(0.068)
=1 if loss>2,000CZK on day 3	0.015*	0.020*	0.015*	0.020*	0.015	-0.023	0.015	-0.026
	(0.008)	(0.011)	(0.008)	(0.011)	(0.041)	(0.068)	(0.041)	(0.068)
=1 if netwin>2,000CZK on day 1	0.002	0.003	0.002	0.004	0.057*	0.077	0.058*	0.082
	(0.008)	(0.011)	(0.008)	(0.011)	(0.033)	(0.053)	(0.033)	(0.053)
=1 if netwin>2,000CZK on day 2	0.002	0.001	0.001	0.001	0.007	-0.017	0.009	-0.014
	(0.008)	(0.012)	(0.008)	(0.012)	(0.031)	(0.048)	(0.031)	(0.048)
=1 if netwin>2,000CZK on day 3	-0.015*	-0.016	-0.015*	-0.016	0.023	0.029	0.023	0.028
	(0.008)	(0.012)	(0.008)	(0.012)	(0.028)	(0.045)	(0.028)	(0.045)
=1 if A or R bet on day 1			-0.009***	-0.012***			-0.007	-0.015**
			(0.002)	(0.003)			(0.005)	(0.006)
=1 if A or R bet on day 2			0.006**	0.007**			-0.004	-0.006
			(0.002)	(0.003)			(0.005)	(0.006)
=1 if A or R bet on day 3			-0.001	-0.000			0.003	-0.005
			(0.002)	(0.003)			(0.005)	(0.006)
=1 if a man, female is the base category	0.052***	0.053***	0.052***	0.053***	-0.018***	-0.007	-0.018***	-0.008
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.007)	(0.005)	(0.007)
R^2, adjusted	0.041	0.036	0.041	0.036	0.020	0.014	0.020	0.014
N (Observations)	76839	55973	76839	55973	29300	21586	29300	21586

Table 8. The betting intensity in sports

The dependent variables are proxies for betting intensity in sports in logarithmic transformation. *Mean_odds* and *mean_stake* stand for the means of daily odds and stakes, respectively. Variable *num_bets* corresponds to the total daily bets computed as the total length of all tickets submitted during the day. The data has a panel structure. The analyzed sample consists of (new) users starting after April 1, 2005. For each user, we excluded the three days at the beginning and the end of his betting history, i.e., $4 \leq \text{day} \leq \text{end-4}$, to filter out early quitters and to minimize end effects. Age group (10 years) fixed effect, and year FE were used as controls. The coefficient estimates for age and year fixed effects are available upon request or on the Internet Appendix. Columns marked as Full contain the results for the entire sample of new users. In the other columns, we exclude bettors with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Heteroskedasticity consistent standard errors are on brackets.

Variables	SPORT					
	Log(mean_odds)		Log(mean_stake)		Log(num_bets)	
	Full	<200	Full	<200	Full	<200
=1 if the (net) win on day 1	-0.716***	-0.692***	0.098***	0.222***	-0.198***	-0.194***
	(0.002)	(0.005)	(0.001)	(0.003)	(0.001)	(0.002)
=1 if near-missed on day 1	-0.084***	-0.043***	-0.053***	-0.070***	0.126***	0.118***
	(0.002)	(0.006)	(0.001)	(0.003)	(0.001)	(0.003)
=1 if loss>2,000CZK on day 1	-0.532***	-0.722***	1.381***	1.817***	-0.235***	-0.419***
	(0.007)	(0.017)	(0.003)	(0.010)	(0.003)	(0.009)
=1 if netwin>2,000CZK on day 1	0.351***	0.178***	0.933***	1.233***	-0.016***	-0.107***
	(0.007)	(0.016)	(0.003)	(0.009)	(0.003)	(0.008)
=1 if the (net) win on day (t-1)	-1.037***	-1.022***	0.255***	0.303***	-0.109***	-0.176***
	(0.002)	(0.005)	(0.001)	(0.003)	(0.001)	(0.002)
=1 if near-missed on day (t-1)	-0.173***	-0.149***	-0.055***	-0.054***	0.254***	0.175***
	(0.002)	(0.006)	(0.001)	(0.003)	(0.001)	(0.003)
=1 if loss>2,000CZK on day (t-1)	-0.750***	-0.998***	2.170***	2.582***	0.012***	-0.416***
	(0.006)	(0.015)	(0.003)	(0.009)	(0.003)	(0.007)
=1 if netwin>2,000CZK on day (t-1)	0.213***	-0.067***	1.577***	2.060***	0.128***	-0.183***
	(0.006)	(0.016)	(0.003)	(0.010)	(0.003)	(0.008)
=1 if a man	-0.009***	0.019***	0.187***	0.254***	-0.025***	-0.062***
	(0.003)	(0.006)	(0.001)	(0.003)	(0.001)	(0.003)
R^2, adjusted	0.083	0.109	0.203	0.249	0.029	0.040
N (Observations)	5,555,229	877,193	7,340,762	1,108,145	7,340,762	1,108,145

Table 9. The betting intensity in lotteries

The dependent variables are proxies for betting intensity in sports in logarithmic transformation. *Mean_odds* and *mean_stake* stand for the means of daily odds and stakes, respectively. Variable *num_bets* corresponds to the total daily bets computed as the total length of all tickets submitted during the day. The data has a panel structure. The analyzed sample consists of (new) users starting after April 1, 2005. For each user, we excluded the three days at the beginning and the end of his betting history, i.e., $4 \leq \text{day} \leq \text{end-4}$, to filter out early quitters and to minimize end effects. Age group (10 years) fixed effect, and year FE were used as controls. The coefficient estimates for age and year fixed effects are available upon request or on the Internet Appendix. Columns marked as Full contain the results for the entire sample of new users. In the other columns, we exclude bettors with more than 200 bets per day (i.e., excluding users employing scripts or bots or potential small bookies). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Heteroskedasticity consistent standard errors are on brackets.

Variables	LOTTERY					
	Log(mean_odds)		Log(mean_stake)		Log(num_bets)	
	Full	<200	Full	<200	Full	<200
=1 if the (net) win on day 1	-0.730***	-1.332***	0.061***	0.347***	0.015***	-0.017**
	(0.008)	(0.043)	(0.002)	(0.008)	(0.003)	(0.009)
=1 if near-missed on day 1	-2.263***	-0.313	0.230***	0.325	0.429***	1.652***
	(0.438)	(1.663)	(0.078)	(0.464)	(0.111)	(0.486)
=1 if loss>2,000CZK on day 1	-0.035	-2.761***	0.814***	2.135***	0.478***	-0.630***
	(0.033)	(0.146)	(0.008)	(0.041)	(0.012)	(0.043)
=1 if netwin>2,000CZK on day 1	1.072***	1.064***	0.135***	0.203***	0.633***	0.127***
	(0.021)	(0.163)	(0.006)	(0.037)	(0.009)	(0.039)
=1 if the (net) win on day (t-1)	-0.989***	-1.942***	0.196***	0.346***	0.264***	0.118***
	(0.007)	(0.036)	(0.002)	(0.007)	(0.003)	(0.007)
=1 if near-missed on day (t-1)	0.419***	-2.390	1.124***	1.533***	0.238***	2.360***
	(0.151)	(1.662)	(0.035)	(0.138)	(0.050)	(0.145)
=1 if loss>2,000CZK on day (t-1)	-0.621***	-2.893***	1.688***	3.101***	1.384***	-0.367***
	(0.022)	(0.137)	(0.007)	(0.040)	(0.010)	(0.042)
=1 if netwin>2,000CZK on day (t-1)	0.959***	0.672***	0.336***	0.780***	0.732***	0.227***
	(0.020)	(0.145)	(0.006)	(0.033)	(0.008)	(0.034)
=1 if a man	0.047***	0.147***	0.246***	0.142***	-0.306***	-0.082***
	(0.004)	(0.021)	(0.001)	(0.005)	(0.002)	(0.005)
R^2, adjusted	0.032	0.063	0.077	0.089	0.052	0.043
N (Observations)	1,482,899	101,655	2,581,850	237,838	2,581,850	237,838