High Bets for the Win?
The Role of Stake Size in Sports Betting*

Radek Janhuba† & Jakub Mikulka‡ (CERGE-EI)§

June 4, 2018

Latest version: http://www.radekjanhuba.com/files/RJJM_Stakes.pdf

Abstract

We analyze the role of stake size in the sports betting market. Our main research question is whether the size of the stake predicts the betting outcomes, i.e. whether bettors can consistently select relatively more profitable events at the most important times. The study utilizes a unique sports betting dataset that includes over 28 million bets by registered customers. We find that bettors are successfully able to vary the stakes in order to increase the probability of their bets winning, but not so much as to increase the net revenue of their bets. The results further suggest that only the most skilled bettors are successfully able to vary the stake size to increase the net revenue. The results are valid regardless of whether bettor fixed effects are included in the analysis, indicating that the relationship between stake and betting outcomes is driven by variation in individual bets.

JEL codes: D12, D81, Z23

Keywords: Betting, Stake, Success, Accumulator Bet

*We would like to thank Randall Filer, Jan Hanousek, Brad Humphreys, Stepan Jurajda, Jakub Steiner, participants in the 2017 YEM Brno and 2017 ESEA Paderborn conferences, and participants in a WVU Brownbag seminar for helpful comments and suggestions. All remaining errors are our own. This study was supported with institutional support RVO 67985998 from the Czech Academy of Sciences.

†Email: radek.janhuba@cerge-ei.cz

‡Email: jakub.mikulka@cerge-ei.cz

§CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic.

1 Introduction

This study examines the role of stake size in the sports betting market. In particular, we are interested in the relationship between stake size and betting outcomes, i.e. the probability of success and bettors' net revenue. The study utilizes a unique sports betting dataset that includes over 28 million bets by customers registered at a large bookmaking company in the Czech Republic. Observing the actual betting transactions rather than only price information is a major advantage of this dataset.

Because we analyze odds set by a bookmaker which are fixed at the time of bet submission and valid for all clients regardless of the stake size, the probability of the bet winning should be unrelated to the stake size. However, it may be possible that bettors place higher stakes in situations in which they believe their bets have a higher probability of success. If they can, on average, vary the stake size successfully, it would indicate that their probabilistic beliefs are important in the decision of the stake size, and that they are, on average, able to correctly identify the profitable betting opportunities.

We also examine the effects of stake size on bettor's net revenue from placing the bet. Because most of the bettors are net losers on the sports betting market, this analysis indicates whether the effects of stake size on the probability of winning are sufficiently strong for the bettors to lose less money.

Note that the examination of the role of stake size inevitably means that we relax the assumption of constant stake size, which has been used by a majority of early literature on sports betting. In this sense, our study extends the analysis of Kopriva (2015) who finds that the assumption of stake size is often violated in the sports betting market.[1]

In theory, investors in financial markets should buy more of those assets which they expect to perform better, but they should buy a limited amount of them as far as they are risk averse (Franklin, Morris & Shin, 2006).[2] In our setting, the stake size can be viewed as the demand for the financial asset in the form of the specific bet. Because the stake size is jointly determined by public information and private beliefs of each customer, our study is connected to the literature on behavioral finance.[3]

Moreover, in experiments focused on the presence of hot-hand and gambler's fallacies in betting behavior, Matarazzo, Carpentieri, Greco & Pizzini (2018) checked the participants' probability estimates in a card drawing game and found that the stake size increased with those probability estimates. This supports the notion that bettors may vary the stake size according to the degree of profitability they perceive in the particular betting opportunity.

Our empirical methodology uses bettors' decisions to spread their bets on multiple opportunities into several single bets or using an accumulator (also known as a *parlay*) bet where multiple opportunities enter the bet. The main distinction between the two types of bets is that in the accumulator bet case, all of the included betting opportunities must win in order to secure any profit. Thus, accumulator bets

[1] See also Feess, Müller & Schumacher (2016) who show that empirical specification controlling for the average stake size leads to more precise estimates of bettors' risk preferences.

[2] Note that while early literature assumed that participation in betting markets is associated with risk-seeking preferences (Sauer, 1998), activity on betting markets is consistent with risk averse, and only seem risk loving due to either skewness seeking preferences (Golec & Tamarkin, 1998) or probability misperceptions (Snowberg & Wolfers, 2010). In fact, most of the recent literature uses CARA utility functions (Feess et al., 2016), which is the functional form assumption used by Franklin et al. (2006).

[3] Given that this study examines sports betting markets, we do not review the behavioral finance literature. A good overview on its topic can be found in Barberis & Thaler (2003). See also Moskowitz (2015).

carry a higher variance of payoffs as well as a lower expected return. Nevertheless, they are extremely popular. In our sample, 91% of betting tickets are accumulators, and only 2.1% users place single bets exclusively.

Because an accumulator bet is effectively a combination of multiple single bets combined together into one betting slip, we denote this combination as a betting *ticket*. Thus, throughout the study, the term *ticket* denotes an accumulator bet, and the term *length* denotes the number of betting opportunities combined into the ticket.

In this study, we do not examine the decisions that lead bettors to select accumulators of given lengths and odds. Rather, we take the characteristics of the placed accumulators as given and we look at whether the stake size, controlling for these observable characteristics of the tickets, predicts the bettors' success rate and net revenue.

The contributions of our study are twofold. First, we show that even in the absence of market regulation regarding the possibility to place single bets, bettors still mostly select accumulator bets rather than single opportunities. While there have been calls to deregulate the sports betting market in the United States, little is known about the possible effects of deregulation. Second, and more importantly, this study is the first to examine the effects of stake size in the context of actually placed betting decisions.

We find that bets with larger stakes are associated with increased success rates, indicating that bettors are able to adjust the stake size in order to win more often. However, due to the overall success rate being relatively small, even though bettors manage to obtain a higher success rate via an increased stake, they do not, on average, lose less money. This is because even if the higher stake translates into a higher chance of a win, the increased net loss from unsuccessful bets cancels out with the increase in net revenue from successful bets.

Further, a segmentation of the bettors based on the value of their fixed effect from the baseline regression, which we take as a proxy of bettors' skill level, reveals that the relationship between stake size and probability of the bet winning is valid for all bettors except for those at the lowest skill level. However, only the most skilled subpopulation of bettors is successfully able to vary the stake size to increase their net revenue from betting, while bettors in the lower half of the skill distribution on average lose more money if they place a higher bet.

Because our sample period includes the moment when the Czech government allowed betting companies to introduce online betting, we further look into the channels through which the specific bets were placed. We find that the effects of a stake on the success rate in bets placed in person at a branch are valid even after the introduction of online betting, but the effect is statistically insignificant for bets placed online. Further, while the general effect of a stake does not generally have statistically significant effects on net revenue, the effect is negative in the case of online bets. Thus, increasing the stake size online leads to a higher net loss.

Our findings are consistent regardless of whether bettor fixed effects enter the regression, indicating that the effects are driven by variation in bets rather than more skilled bettors placing higher bets.

The remainder of this study is structured as follows. Section 2 introduces the basic terminology used in the study and presents characteristics of single and accumulator bets and provides a brief literature review (readers familiar the mechanics of single and accumulator bets may prefer to skip Section 2.1). We describe the dataset in Section 3 and the empirical methodology in Section 4. Section 5 discusses the empirical results, and Section 6 analyzes their robustness. Section 7 concludes.

2 The Betting Market

From the point of view of financial markets, studying betting markets is useful as they usually contain a relatively simple structure and provide well-defined, measurable outcomes. However, betting markets are organized differently from most of other financial markets. The purpose of this section is to introduce a typical European betting market, define the terminology of odds used in this study, and present the single and accumulator types of betting tickets. Readers familiar with the betting markets in continental Europe may prefer to move on to Section 2.2.

We study a *fixed odds* market organized by bookmaking companies. When a bettor wants to place a bet on a selected outcome of a game, he looks at the price information, which is summarized by the *odds*. He then places a *stake*, indicating the amount which the bettor risks. At that moment, the payoffs are defined and, although the odds may change for newly negotiated bets, the odds of the particular bet remain fixed.

Note that throughout the study, we use decimal odds, meaning that if a stake s is placed with odds x, the bettor receives amount sx if the bet wins and 0 otherwise (hence, net revenue is $s(x-1)$). Thus, a fair bet with a chance of 50% would be associated with the odds of 2. These decimal odds are the usual form of odds used in continental Europe, where our dataset comes from, and are also convenient for their mathematical properties (see below).

2.1 Single and Accumulator Bets

A single bet is a bet consisting of one betting opportunity only. Therefore, placing a stake s_1 on an opportunity 1 with odds x_1, and analogically with s_2 and x_2, the outcomes of both opportunities will be evaluated separately. Thus, if bet 1 wins, the bettor receives s_1x_1, regardless of the outcome of bet 2, and vice-versa. However, if a bettor places a stake s on an accumulator bet of the two opportunities with the same odds, the total ticket odds are calculated as x_1x_2, and both outcomes have to be evaluated as winning for the bettor to receive an amount sx_1x_2.

Suppose a bettor wants to place bets on two games, each of which has two possible outcomes with a probability of 0.5, that is, odds of 2. The bettor wants to place a bet of \$100. He can either place two bets of \$50, one on each game, or he can place an accumulator bet, where the odds of each single bet will multiply, obtaining a total odds of 4. However, for the bettor to win anything, both of the bets have to win. If either of them loses, the bet returns 0.

As summarized in Table 2.1, the expected payoff in this simple case is 100 for both types of bets. Therefore, assuming there are two clients which each place a bet on the opposite combination of bets, the betting company will receive bets of 200 and also pay out 200.

In practice, the betting company will often set up an *overround* such that it can reduce the amount paid out. Effectively, this means that the odds on a bet with a 50% chance of success will be set as 1.8. Under this alternative setting, the bettor's

Table 2.1: A Simple Bet Example

Probability	Bet 1	Bet 2	Payout of singles	Payout of accumulator
A. No overround of the betting company (50% bet has odds 2.0)				
0.25	Won	Won	200	400
0.25	Won	Lost	100	0
0.25	Lost	Won	100	0
0.25	Lost	Lost	0	0
			$E(X) = 100$	$E(X) = 100$
B. 10% overround of the betting company (50% bet has odds 1.8)				
0.25	Won	Won	180	324
0.25	Won	Lost	90	0
0.25	Lost	Won	90	0
0.25	Lost	Lost	0	0
			$E(X) = 90$	$E(X) = 81$

expected outcomes will differ. Specifically, while the expected payout of the single bet is equal to 90, the expected payout from an accumulator bet is equal to 81. Therefore, selecting the accumulator bet is not only a matter of higher variance of return, but is also choosing the lower expected payoff.

Formally, assume a constant overround of the betting company δ. Suppose also that the consumer likes N games and wants to bet on them. Each of these bets has odds x_k where k ranges from 1 to N. The consumer wants to spend a total income of M. Moreover, denote π_k the probability of a bet k winning and w_k the payout of a betting ticket if it wins.

In a perfect world with no transaction costs and zero overround of the betting company, the fair odds are set such that $x_k = 1/\pi_k$. Taking the inverse, we obtain $\pi_k = 1/x_k$, which would be the underlying probability based on observing the odds x_k in the perfect world. The expected payoff from placing the selection of matches in the form of single bets is

$$E_{single} = \sum_{k=1}^{N} \pi_k w_k = \sum_{k=1}^{N} \frac{1}{x_k} \frac{M}{N} x_k = M, \tag{1}$$

while the expected payoff from placing an accumulator bet is

$$E_{accum} = \pi_w w = \left(\prod_{k=1}^{N} \frac{1}{x_k} \right) M \left(\prod_{k=1}^{N} x_k \right) = M. \tag{2}$$

Suppose now the betting company charges an overround δ, meaning that it wants to reduce the payout by this percentage. This means that the probability of π_k is no longer an inverse of the odds, but rather $\pi_k = (1 - \delta)/x_k$. The expected return of the two types of bets thus becomes

$$E_{single} = \sum_{k=1}^{N} \pi_k w_k = \sum_{k=1}^{N} \frac{1-\delta}{x_k} \frac{M}{N} x_k = (1 - \delta)\, M, \tag{3}$$

$$E_{accum} = \pi_w w = \left(\prod_{k=1}^{N} \frac{1-\delta}{x_k} \right) M \left(\prod_{k=1}^{N} x_k \right) = (1 - \delta)^N M. \tag{4}$$

As δ must from the definition be between 0 and 1, the expected return of the accumulator bet will always be lower than the one from placing everything on single bets.

2.2 Literature Review

Empirical studies of sports betting markets have largely concentrated on the examination of available odds. Thus, the research has focused on price information rather than actual transactions. As our study concentrates on the analysis of actually placed transactions in the context of accumulator bets, we do not review this literature here. Sauer (1998) provides an overview of the early sports betting literature.

To our knowledge, only a handful of studies have used actual decisions of individual bettors on the market, rather than only the price information available publicly. Gainsbury & Russell (2015) use one year of data from a large Australian betting company to provide descriptive statistics of the composition of bets placed by bettors. They report that most bettors lose money and that there is a substantial variation in the stakes placed.

Using an individual level dataset on New Zealand sports bets, Feess et al. (2014) examine the role of experience in betting behavior. Their results suggest that more experienced bettors tend to bet more on favorites and are able to choose more profitable betting events.

Andrikogiannopoulou & Papakonstantinou (2011) use a random subsample of one-hundred bettors from a large online betting company to examine the extent of the *favorite longshot bias* (FLB), a stylized fact in the betting literature in which bets on favorites tend to produce higher returns than bets on outsiders.[4] Their results confirm the overall existence of the FLB. Further, they claim that the FLB is utilized by about 2% of bettors who are able to achieve a positive net revenue.

A handful of studies have examined the decisions between single and accumulator bets in a theoretical setting. Grant (2013) provides a formal treatment of the two types of bets and strategy types associated with them. He finds that strategies with single bets tend to outperform those with accumulator betting tickets. Generally, the only situation in which the bettor is better off with an accumulator bet is when he is able to consistently outperform the bookmaker in predicting the probability of match outcomes, and hence combine better-than-fair bets into an accumulator.

[4] See Ottaviani & Sørensen (2008) for an overview of the main explanations of the FLB.

Zafiris (2014) provides an example of such betting strategy.

To the best of our knowledge, no study has focused on the empirical relationship between stake size and the success of the bet. This is the gap that our study fills.

3 Data

We use a unique sports betting dataset that was provided by a major Czech betting company. The dataset includes complete betting decisions of the company's registered customers, consisting of tickets and betting opportunities on each ticket. At the ticket level, we observe the betting opportunities selected, length, total odds, time of placing the ticket, and the channel through which it was placed. Further, we see the individual odds for each betting opportunity selected on the ticket. The information related to customers is the region of residence, age, and gender. The time span of the data is January 2005 – February 2012. In total, the dataset includes 112,409 registered clients who placed 28,543,717 tickets.[5]

The dataset reveals that bettors usually decide to place accumulator rather than single bets. Of the total number of 28,543,717 tickets, 25,962,787 (91%) are accumulator tickets and 2,580,930 (9%) are tickets consisting of single bets. Moreover, only 2,413 (2.1%) users place single bets exclusively. These values reveal that accumulator bets form an important part of the market.[6] Arguably, the market is thus comparable to markets in areas where gambling regulation restricts betting to only

[5]Note that the Czech regulations also allow bettors to place combination tickets which effectively combine many single and accumulator tickets into one betting slip. For example, if a bettor selects three opportunities and places them on a combination ticket, the resulting ticket may effectively include a three-way accumulator, three two-way accumulators, and three singles. Due to their rather complicated analysis and specific behavior, we exclude combination tickets from the analysis.

[6]Note that the betting company that provided the data is not in any way specialized in offering or promoting accumulator bets. Thus, the property of accumulator tickets being popular holds for the market in general.

be allowed in the form of accumulator bets on specific betting opportunities, such as e.g. Canada or the US state of Delaware. These percentages also show that even in the absence of regulation of single bets on the betting market, most of the bettors choose to place single bets only occasionally.

The descriptive statistics on the user level are presented in Table 3.1. We can see that there is substantial variation in all of the variables. As expected, most of the distributions are positively skewed, which can be seen from the higher value of the mean as compared to the median. Interestingly, the net revenue is skewed negatively, indicating that there are relatively more people with large net losses rather than those with large wins.

Table 3.1: Descriptive Statistics - User Level

	N	Mean	S.D.	Min	Median	Max
Tickets Bet	112,409	403	1,199	1	58	59,392
Success Rate	112,409	.097	.14	0	.044	1
Net Revenue	112,409	-4,792	48,204	-3,833,882	-400	5,096,647
Stake[A,1]	112,409	158	819	10	52	126,885
Length[A]	112,409	7.4	4	1	6.8	67
Odds[A]	112,409	659	4,089	1.06	35	321,288

[A] Variables marked with A denote averaged values over all tickets placed by the user.
[1] Stake denominated in Czech Korunas.
Source: Authors' calculation

Turning to Table 3.2 showing the descriptive statistics at the ticket level, we can see that there is a considerable variation in all of the variables in question. Because all of the distributions are positively skewed, the median gives a better picture of what the typical ticket looks like. It reveals that a typical ticket has a stake of 40 CZK, has five betting opportunities on it, and has odds of 15.91. A simple calculation reveals that if all of the opportunities were the same, each would have an odds of approximately 1.74, meaning a chance of success of just over 50%.

Compounding this to take into effect the length of 5, the probability of a median ticket winning comes to about 3.7%. Note that this is much lower than the implied probability of the median odds ticket of 5.7%. This difference is attributed to the overround being compounded in favor of the bookmaker (see equation 4 in Section 2.1).

Table 3.2: Descriptive Statistics - Ticket Level

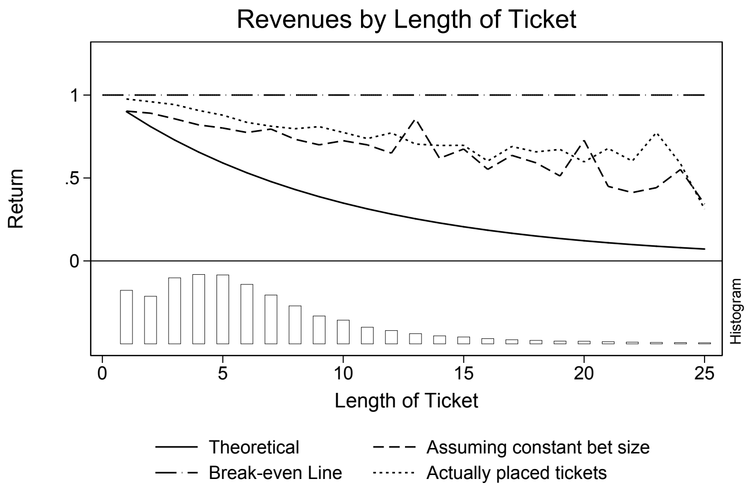
	N	Mean	S.D.	Min	Median	Max
Ticket Success	28,543,717	0.11	0.31	0	0	1
Net Revenue	28,543,717	-18.87	1630.94	-397,090	-30	1,082,115
Stake	28,543,717	150.52	865.22	10	40	500,000
Length	28,543,717	6.61	5.47	1	5	99
Odds	28,543,717	835.01	9466.41	1.01	15.91	652,845

Source: Authors' calculation

The relationship between the length of a ticket and its expected profitability is depicted in Figure 1. The solid line is a theoretical profitability line assuming a constant overround of 10%. The horizontal line at 1 is the break-even line. With the exception of two specific lengths (which are caused by extremely improbable wins of large magnitudes), all ticket positions are between the two lines, suggesting that bettors are on average not profitable, but also perform better on average than what is implied by the theoretical calculation of revenues from accumulator bets.

More interestingly, the dashed line shows the hypothetical profitability of tickets under the assumption of a constant bet size (i.e. it calculates the fraction of the actually placed tickets that won) while the dotted line also takes into account the stakes. The dotted line lies above the dashed line for all tickets until the length of approximately 25, which may be interpreted as an indicator of bettors utilizing the variation in bet size in order to increase their profitability, and is, thus, a direct

Figure 1: Length and Profitability of Accumulator Bets



motivation for testing our research question, i.e. the effect of the stake size on the probability of success and net revenues.

4 Methodology

We estimate a basic fixed effect model to analyze the effects of the ticket characteristics on the betting outcomes. The model follows the equation

$$y_{itk} = \beta_0 + \beta X_{itk} + \gamma C_{itk} + \alpha_i + \theta_t + \varepsilon_{itk} \tag{5}$$

where the dependent variable y_{itk} is either a dummy variable for whether the particular ticket won, or the net revenue of the ticket, and the subscript itk may be read as "ticket k placed by bettor i at time t". The vector X_{itk} includes the ticket's stake, length (number of betting opportunities), and inverse odds of the ticket to control for the overall riskiness of the selected games. These inverse odds serve as a proxy for the ticket's general probability of winning.

The vector C_{itk} includes control variables which could potentially be associated with the probability of the ticket's win. Specifically, we include dummies for specific channels through which the bet can be placed. Specifically, the two variables are Internet for bets placed online and telephone for bets placed via phone.

The coefficient α_i is a client-specific fixed effect and thus controls for observed and unobserved time-invariant factors influencing the probability of a win for each ticket, for instance the bettors' time-invariant preferences and skills. The possibility to include this fixed effect is one of the major advantages of our dataset. We also include a weekly fixed effect θ_t which captures the time variation in the overall success rate.[7] The combination of time and user fixed effects controls for the possible presence of behavioral biases that were previously found in sports betting markets (e.g. the home bias documented by Braun & Kvasnicka (2013) and Staněk (2017)).

The chosen methodology is relatively robust against the implications of the favorite longshot bias. In our analysis, the FLB implies that the betting opportunities with lower odds tend to have lower overround in comparison with higher-odds bets. Therefore, in the case of the two accumulator tickets with the same odds but different lengths, it is not clear which of them has higher winning probability, as the length disadvantage is compensated by lower over-round for the small odds opportunities. Nevertheless, because the effect of the coefficient on stake has to be interpreted ceteris paribus, controlling for the ticket's length and odds eliminates the methodological issues posed by the FLB.

Due to a relatively high number of observations, we use a linear probability model in all binary outcome regressions. This allows us to directly compare the results of success and net revenue regressions and to avoid the issues with the estimation of

[7] It is likely that in the case of some specific events, such as the ice hockey World Championships, there may be a systematic pattern in stake size and the success rate, which would mainly be determined by the Czech national team.

large-scale limited dependent variable models with fixed effects.

5 Results

5.1 Baseline Analysis

The results of the baseline analysis are shown in Table 5.1. Panel A presents the results of a linear probability model for success where the left hand side variable takes the value of 1 if the ticket won and 0 otherwise. The coefficient of the stake, our variable of interest, is positive, which suggests that the tickets with a higher stake win relatively more often compared to similar tickets with a lower stake. This indicates that bettors place higher stakes on tickets which they value more and these in turn have a higher realized chance of winning.

Table 5.1: Baseline Results

	Panel A: Success				Panel B: Net Revenue			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000244***	.000234***	.000094***	.000095***	3.0629	3.095	1.7589	1.7596
	(2.6e-05)	(2.5e-05)	(1.6e-05)	(1.6e-05)	(2.9)	(2.93)	(3.97)	(3.97)
Length	-.000053***	-.00005***	-.00006***	-.000071***	-.3535***	-.35964***	-.20052**	-.20449**
	(1.8e-05)	(1.8e-05)	(2.2e-05)	(2.2e-05)	(.084)	(.082)	(.101)	(.103)
Inv. Odds	.93169***	.93168***	.92757***	.92741***	-31.165	-34.677	-38.849	-38.671
	(2.1e-03)	(2.2e-03)	(2.4e-03)	(2.4e-03)	(22.7)	(22.4)	(33.4)	(33.4)
Internet		-.000589**	-.002629***	-.000213		5.9524***	2.9831	7.9086***
		(2.5e-04)	(3.7e-04)	(3.9e-04)		(2.03)	(2.83)	(2.9)
Phone		.003603***	-.002619*	-.002271		-6.8595	22.102	22.878
		(8.8e-04)	(1.4e-03)	(1.4e-03)		(17.3)	(21.3)	(21.5)
Constant	-.007419***	-.007357***	-.005803***	.002166	-17.226***	-18.298***	-16.793***	-7.4918
	(2.8e-04)	(2.8e-04)	(3.9e-04)	(1.5e-03)	(1.67)	(1.83)	(2.64)	(8.03)
User FE			Yes	Yes			Yes	Yes
Week FE				Yes				Yes
N	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717

OLS results. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

The results of the regression with the ticket's net revenue as the dependent variable are presented in Panel B. Note that the results are very different compared to the results in Panel A, as the stake coefficient is not significantly different from zero at any conventional level of confidence. This implies an interesting result: while bettors can influence the rate at which their tickets win using variation in stake size, they cannot do so to such an extent that they would lose less money. As the baseline success rate is generally quite low, the higher success rate does not help bettors to increase net revenue, mainly because placing a higher stake most of the time leads to a higher loss.[8]

The results are qualitatively very similar regardless of whether user fixed effects enter the estimation, which suggests that the relationship between stakes and probability of a win is driven mostly by the variation within individual bets rather than the variation between the bettors.

In terms of magnitude, the baseline effect means that if the stake increases by 100 Czech Korunas, the probability of the ticket's success increases by 0.000095. Given that the unconditional success rate of an average ticket is 0.11, this would correspond to an increase of about 1% of the average success rate. Thus, while the effect is small, it may arguably be viewed as non-negligible.

The results also reveal that the longer the ticket, the lower the chance of a win, which is consistent with the theoretical reasoning outlined in Section 2.1, although the size of the coefficient is quite low. The small size of the effect may be connected to the fact that longer tickets tend to be filled with lower-odds events, which can be

[8]Imagine a situation in which a client places a stake of $10 on ticket A and a stake of $100 on ticket B, with both tickets having odds of 5. In the success regressions, the dependent variable will be 0 or 1 in both cases. In the net revenue regressions, the dependent variable will be -10 or 40 in the case of ticket A, but -100 and 400 in the case of ticket B. Thus, if the overall success rate is relatively small, its increase after placing the higher stake cancels out with the increased net loss most of the time.

more profitable to bet due to the favorite longshot bias as discussed in Section 4. The analysis also reveals that the higher the odds, the lower the chance of winning, a result that is expected as odds are proportional to the inverse of the probability of the particular match winning.

Although columns (2) and (3) may suggest that bets placed on the Internet and telephone have a different general probability of success, this result disappears after controlling for time fixed effect. However, as revealed in column (8), bets placed over the Internet generally have a higher net revenue as compared to bets placed in person at a branch. We return to the discussion on the role of channels through which the tickets are placed in Section 5.3.

5.2 Role of Skills

This section looks at the possible heterogeneity of the effect based on the betting skills of the customer. We expect that the most skillful bettors should be able to better identify the least profitable betting opportunities and utilize the variation in stakes to improve the betting outcomes.

However, it is difficult to assess the skills of the individual bettors. We choose to utilize the value of the individual fixed effect in the baseline success regression to proxy for bettor skill, and segment bettors into four categories based on the value of this fixed effect. While this is not an ideal solution, we feel that the fixed effect is the best available measure of one's (unobservable) skill. Alternative measures based on experience with betting are discussed in Section 6.2.

The results are presented in Table 5.2. In Panel A, the main result of the stake being positively correlated with the success of individual bets is valid for all of the

Table 5.2: Role of Skills

	Panel A: Success				Panel B: Net Revenue			
Skills	Lowest (1)	Low (2)	High (3)	Highest (4)	Lowest (5)	Low (6)	High (7)	Highest (8)
Stake/100	-9.8e-06	.000057**	.000079***	.000062**	-54.313***	-13.909**	.97661	15.914***
	(5.2e-05)	(2.8e-05)	(2.3e-05)	(2.5e-05)	(9.14)	(6.17)	(7.14)	(5.86)
Length	-.000573	-.000034	.000013	-.000822***	.92145**	-.25734**	-.29521**	-.38181
	(5.3e-04)	(3.2e-05)	(1.2e-05)	(5.6e-05)	(.4)	(.127)	(.142)	(.352)
Inv. Odds	.63963***	.85877***	.94023***	1.0503***	147.05***	34.08	-43.852	-86.348
	(.03)	(2.2e-03)	(1.2e-03)	(2.4e-03)	(40.4)	(41.4)	(63.8)	(58)
Internet	.004677	.001511*	-.0011***	.000741	-30.49	3.5865	7.4733*	24.57**
	(5.4e-03)	(8.1e-04)	(3.9e-04)	(1.5e-03)	(18.6)	(2.93)	(4.27)	(10.4)
Phone	.016051	.000767	-.003564*	-.002406	28.939	-22.295	-3.7155	45.587
	(1.0e-02)	(2.7e-03)	(1.8e-03)	(3.3e-03)	(60.1)	(19.7)	(11.8)	(68)
Constant	-.004009	-.004134	.002581*	.013157***	-5.6576	-12.692	4.9797	-26.506
	(9.5e-03)	(3.5e-03)	(1.5e-03)	(4.7e-03)	(21.5)	(10.6)	(10.4)	(39)
N	709,499	6,612,524	17,520,693	3,701,001	709,499	6,612,524	17,520,693	3,701,001

OLS results. Skill categories broken down based on the value of the fixed effect from baseline regression. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

bettors except for those with the lowest skills.[9] Even though the sample is divided into groups of the same number of bettors, bettors with the lowest skills also place the fewest tickets. This may indicate that unsuccessful bettors quit or reduce their betting frequency when their tickets do not win.

Panel B shows the results in the net revenue regressions. The results reveal that at least half of the bettors actually lose more money by placing a higher bet. At the same time, bettors at the highest skill level can successfully vary the stake size in order to increase their net revenue. Overall, the regressions in Panel B indicate that the higher the skill level of the bettor is, the more successful that bettor is in varying the stake size in order to increase his net revenue.

[9] Note that in terms of financial terminology, these bettors might be labeled as noise traders.

5.3 Channels

There are three possible channels bettors may use to place bets. First, they may visit the betting branch and place their bet in person. Second, they may place their bets via a phone call. Third, beginning from 2009, they may bet online on the website of the betting company.

The betting behavior may differ depending on which channel is used, for instance, due to a difference in the degree of easiness to buy multiple tickets at any time over the Internet, or the availability of sport-related information. Further, it is possible that the subpopulation of online bettors may be different from the subpopulation using mainly the betting branches.

All of these differences translate to differences in typical bets placed, which are presented in Table 5.3. To have a better comparison of online and offline bets for the same periods, we split the bets placed at branches into two subperiods (we do not split the phone bets beacause there is a relatively small share of them). The comparison shows that the online bets are comparable to bets placed through other channels in terms of average ticket revenue and average stake. However, online tickets are on average shorter and with higher odds than tickets placed at branches.

The telephone tickets are on average shorter and with a higher average stake and lower average odds in comparison with all other channels. The comparison between the two subperiods for the tickets placed at branches reveals that most of the characteristics remain similar except for rising average odds, decreasing ticket revenue, and success over time.

To see whether the effect of stake size on the success rate and net revenues differ

Table 5.3: Descriptive Statistics - Ticket Level by Channel

	Branch 2005-2008	Branch 2009-2012	Internet	Telephone
Ticket Success	0.0868	0.0759	0.162	0.230
	(0.282)	(0.265)	(0.368)	(0.421)
Net Revenue	-18.43	-23.70	-16.70	-8.636
	(1611.3)	(1213.4)	(968.6)	(5519.9)
Stake	136.9	115.1	129.9	858.8
	(740.7)	(493.2)	(610.4)	(3352.8)
Length	7.091	7.070	5.724	3.964
	(5.221)	(5.395)	(5.890)	(3.218)
Odds	440.9	920.2	1469.6	54.56
	(4232.0)	(10419.8)	(14004.8)	(1266.2)
N	12,954,913	6,606,718	8,164,756	817,330

Mean values (standard deviations in parentheses). Source: Authors' calculation

for online and telephone bets, we segment tickets by the channel they were placed through and estimate the analysis separately for each segment. The results are presented in Table 5.4.

In the case of success regressions shown in Panel A, the stake size significantly positively affects the success of the ticket in both subperiods of betting at branches and by phone, while the effect is positive but not statistically significant for online bets. This non-existence of the effect significance may be explained by the fact that each regression coefficient presents the effect of increasing the stake size compared to other tickets placed through the particular channel. Because, as shown in Table 5.3, online bets have a higher overall success rate, it is possible that the effect is not as pronounced in their case. Moreover, note that the mean odds of Internet-based tickets are much higher than in the case of other channels. As shown below in Section 5.4, the effect is not found for tickets with extremely high odds, which may be explained by a lower variation in the stake placed on such tickets.

Table 5.4: Decomposition by Betting Channels

	Panel A: Success				Panel B: Net Revenue			
Channel	Branch (1)	Branch (2)	Internet (3)	Phone (4)	Branch (5)	Branch (6)	Internet (7)	Phone (8)
Stake/100	.000127***	.000187***	.000026	.000078***	-3.4112	3.0635	-13.829***	13.822
	(2.8e-05)	(5.8e-05)	(3.4e-05)	(2.8e-05)	(4.51)	(9.99)	(2.45)	(8.99)
Length	-.000144***	-.000208***	.000044**	-.00027	-.40889**	-.27425	-.21382**	2.2582
	(4.5e-05)	(4.3e-05)	(2.2e-05)	(2.1e-04)	(.181)	(.189)	(.09)	(3.01)
Inv. Odds	.9212***	.88499***	.94637***	.95751***	5.3456	-89.348	64.398***	-243.61
	(5.6e-03)	(5.7e-03)	(1.6e-03)	(4.6e-03)	(49.9)	(89.1)	(13.1)	(204)
Constant	.002796*	-.010519***	-.017267***	-.00975	-4.3852	-22.144***	-.39022	-114.66
	(1.7e-03)	(1.5e-03)	(5.4e-03)	(8.2e-03)	(4.6)	(7.14)	(10.2)	(138)
Beginning	2005	2009	2009	2005	2005	2009	2009	2005
End	2008	2012	2012	2012	2008	2012	2012	2012
N	12,954,913	6,606,718	8,164,756	817,330	12,954,913	6,606,718	8,164,756	817,330

OLS results. Sample broken by betting channels, with branch broken according to pre and post Internet periods. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

Panel B shows the results for the net revenues. Similarly to success regressions,the results for bets over the Internet are different as compared to other betting channels. While in the segments of branch and phone bets the positive coefficient of stake size is insignificant, in the segment of online bets the coefficient of interest is significantly negative. The likely reason is that, because the increased stake is not associated with a higher probability of the ticket winning, and the overall success rate is very small, placing a higher bet over the Internet generally means that a bettor loses more money.

Further, the comparison of tickets placed at branches in the two periods (2005-2008 and 2009-2012) reveals that the results remain qualitatively the same. Although neither of the effects is statistically significant, the results present weak evidence for the effect of a stake rising over time.

5.4 Relationship with Ticket Length and Odds

In this section, we examine whether our baseline results differ for different levels of ticket length and ticket odds. Table 5.5 shows the baseline model estimated for the four subsamples given by the quartiles of tickets ordered by their length. The results suggest that the positive effect of a stake on the probability of success is valid for all but the longest tickets. The effect on the longest category of tickets is also positive, although not statistically significant. This may be explained by the low variation in stake sizes for betting tickets with eight and more opportunities, and such tickets usually having a very low success rate. Thus, this segment of tickets has a low variation in both the dependent and main explanatory variables, making the statistical inference more difficult.

Table 5.5: Ticket Segmentation by Length

	Panel A: Success				Panel B: Net Revenue			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000083***	.000145***	.000061*	.00007	7.4495	-1.1957	-12.062**	-13.364
	(2.0e-05)	(3.3e-05)	(3.6e-05)	(4.6e-05)	(6.17)	(4.34)	(4.69)	(14.6)
Length	-.002901***	-.002578***	-.00154***	-.000107***	-2.4922	-.80709	-1.9146***	-.60629***
	(6.4e-04)	(3.3e-04)	(1.2e-04)	(1.0e-05)	(1.9)	(1.36)	(.577)	(.136)
Inv. Odds	.97212***	.90124***	.85157***	.77344***	-53.028	-14.55	25.876	-45.139
	(2.0e-03)	(7.2e-03)	(2.8e-03)	(3.8e-03)	(52.4)	(37)	(40.3)	(132)
Internet	-.002793**	-.000214	-.000362	-.000119	12.597	10.223***	1.8769	1.6807
	(1.2e-03)	(8.5e-04)	(5.2e-04)	(3.6e-04)	(8.9)	(3.78)	(2.65)	(4.33)
Phone	-.008652***	-.001519	.001097	-.000631	-20.457	43.218**	95.186*	-39.709
	(2.7e-03)	(2.5e-03)	(2.6e-03)	(2.4e-03)	(36.1)	(20.1)	(52.6)	(35.8)
Constant	.008317	.022939***	.01509***	.002589**	15.407	-2.528	20.714*	-4.8197
	(5.4e-03)	(3.8e-03)	(2.3e-03)	(1.2e-03)	(25.5)	(18.2)	(11.8)	(8.06)
Min. Length	1	3	5	8	1	3	5	8
Max. Length	2	4	7	99	2	4	7	99
N	4,876,263	6,528,883	8,548,318	8,590,253	4,876,263	6,528,883	8,548,318	8,590,253

OLS results. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

Table 5.6: Ticket Segmentation by Odds

	Panel A: Success				Panel B: Net Revenue			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000077***	.000064**	.000113**	.000059	2.8582	-8.457	16.405	.93127
	(1.9e-05)	(2.7e-05)	(4.5e-05)	(5.5e-05)	(4.81)	(5.29)	(25)	(36.7)
Length	-.004845***	-.001421***	-.000469***	-.000037***	-3.2528***	-1.4703***	-1.1174***	-.49075***
	(1.8e-04)	(4.9e-05)	(1.9e-05)	(3.6e-06)	(.607)	(.258)	(.329)	(.098)
Inv. Odds	.95946***	.86766***	.77373***	.68343***	-21.514	-.22465	-292.23	-371.42
	(4.6e-03)	(3.7e-03)	(5.3e-03)	(6.9e-03)	(39.6)	(49.4)	(237)	(686)
Internet	.000064	-.000306	.000246	-.000129	15.634**	5.8514*	8.4315	2.4721
	(1.3e-03)	(6.2e-04)	(3.6e-04)	(1.5e-04)	(7.15)	(3.3)	(5.34)	(4.71)
Phone	-.005151**	-.00143	.001276	.00067	41.273	13.661	-41.365	132.24
	(2.5e-03)	(2.2e-03)	(1.7e-03)	(1.0e-03)	(28.3)	(24.6)	(38.7)	(210)
Constant	.02214***	.00838***	.008238***	.0022***	7.1649	-3.7103	3.4587	-5.5317
	(4.8e-03)	(2.6e-03)	(1.5e-03)	(6.6e-04)	(29.8)	(9.06)	(8.55)	(10.7)
Min. Odds	1.01	5.51	15.91	58.54	1.01	5.51	15.91	58.54
Max. Odds	5.5	15.9	58.53	652845	5.5	15.9	58.53	652845
N	7,135,427	7,136,274	7,136,019	7,135,997	7,135,427	7,136,274	7,136,019	7,135,997

OLS results. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

Table 5.6 shows similar subsampling on quartiles, but by odds. Similarly to decomposition based on ticket length, the results reveal that for all subsamples the effect of a stake on a ticket's success is relatively homogeneous except for the fact that we lose significance for the highest quartile of odds, which may be caused by the relatively low variance in stakes for the highest quartile as customers tend to bet lower amounts on high odds tickets.

6 Robustness Checks

6.1 Non-Linear Relationship

In this section, we report results on the baseline regression allowing for nonlinearity of the effects of stake size on betting outcomes. The results, including a squared value of a stake, are presented in Table 6.1. The results imply that, even though the second power of stake is statistically significant, its value is so low that the resulting effect of a stake is almost a straight line. Hence, we abstain from using the second power in the main results of the study.

The coefficients of all of the control variables in the model are qualitatively comparable to their respective values in the baseline estimation presented in Table 5.1.

Table 6.1: Robustness for Non-Linear Effect

	Panel A: Success				Panel B: Net Revenue			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000339***	.000326***	.000134***	.000135***	-.34283	-.35879	-3.5883	-3.5889
	(3.0e-05)	(2.9e-05)	(2.0e-05)	(1.9e-05)	(3.08)	(3.12)	(3.82)	(3.82)
$(\text{Stake}/100)^2$	-1.3e-07***	-1.2e-07***	-4.4e-08***	-4.3e-08***	.004618	.004624	.005833	.005833
	(2.5e-08)	(2.4e-08)	(1.4e-08)	(1.4e-08)	(6.1e-03)	(6.2e-03)	(6.5e-03)	(6.5e-03)
Length	-.000052***	-.00005***	-.000059***	-.000071***	-.3826***	-.37552***	-.23551**	-.24334**
	(1.8e-05)	(1.8e-05)	(2.2e-05)	(2.2e-05)	(.083)	(.08)	(.099)	(.1)
Inv. Odds	.93101***	.93104***	.92726***	.9271***	-6.7529	-10.73	3.0129	3.1291
	(2.1e-03)	(2.2e-03)	(2.4e-03)	(2.4e-03)	(21.3)	(21.4)	(30.1)	(30.1)
Internet		-.000535**	-.002604***	-.000189		3.942**	-.35535	4.6227*
		(2.5e-04)	(3.7e-04)	(3.9e-04)		(1.9)	(2.68)	(2.72)
Phone		.003171***	-.0027*	-.002352		9.3195	32.918*	33.774*
		(8.6e-04)	(1.4e-03)	(1.4e-03)		(15.2)	(19.6)	(19.7)
Constant	-.007471***	-.007411***	-.005827***	.002146	-15.336***	-16.253***	-13.586***	-4.8221
	(2.8e-04)	(2.8e-04)	(3.9e-04)	(1.5e-03)	(1.81)	(1.98)	(2.61)	(7.61)
User FE			Yes	Yes			Yes	Yes
Week FE				Yes				Yes
N	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717

OLS results. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

6.2 Alternative Measures of Skill

In this section, we provide the results of the examination of the role of skills via an alternative measure through bettors' experience. Specifically, we segment bettors based on their number of tickets bet and the number of days they were active on the betting market. However, note that the distribution of bettors is not uniform, but rather skewed with most bettors betting a relatively low number of tickets and very few of them betting high numbers of tickets. Hence, to have the total number of bets, and thus the number of observations, more or less equal in all segments, the segmentation is based on the 99th, 95th, and 85th quantiles of the distribution of bettors according to their total number of tickets placed.

The results for both segmentations are presented in Tables 6.2 and 6.3. The results on success rates, shown in Panel A of both tables, reveal a similar pattern as the skill measure based on the value of the fixed effect from the baseline regression. In all cases, the results are valid for all bettors except those with the relatively lowest skills.

Table 6.2: Customer Segmentation by Number of Tickets Placed

	Panel A: Success				Panel B: Net Revenue			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000011	.00009***	.000161***	.000267***	1.4644	4.5715	1.8699	-8.247**
	(2.2e-05)	(2.8e-05)	(3.5e-05)	(7.4e-05)	(8.48)	(7.5)	(3.2)	(3.28)
Length	.000052**	-.00003	-.00007**	-.00021***	-.017147	-.15705	-.19326	-.50197**
	(2.3e-05)	(2.0e-05)	(3.2e-05)	(7.2e-05)	(.136)	(.113)	(.211)	(.245)
Inv. Odds	.94089***	.93482***	.92389***	.90456***	-25.047	-60.424	-43.952	26.874
	(1.6e-03)	(1.7e-03)	(3.4e-03)	(.011)	(60.6)	(68.5)	(32.7)	(24.1)
Internet	-.003737***	-.000149	.000799	.000511	3.7019	11.242	12.648***	.066719
	(7.6e-04)	(6.6e-04)	(7.2e-04)	(8.6e-04)	(7.52)	(6.87)	(3.34)	(2.69)
Phone	-.003288	-.001814	-.004424*	.003468	97.808	-2.2532	3.6893	14.225
	(2.9e-03)	(2.4e-03)	(2.4e-03)	(5.3e-03)	(94.3)	(28)	(13.9)	(19)
Constant	.006668***	-.000282	.001552	.002535	9.2845	-19.2	-15.121	-2.1387
	(2.4e-03)	(2.3e-03)	(2.4e-03)	(5.1e-03)	(20.2)	(13.3)	(15.9)	(7.8)
N	6,051,992	7,318,062	8,446,096	6,727,567	6,051,992	7,318,062	8,446,096	6,727,567

OLS results. Sample broken by the total number of tickets placed by the customer. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

The results for net revenue reveal a slightly different pattern. Specifically, there

Table 6.3: Customer Segmentation by Active Days

	Panel A: Success				Panel B: Net Revenue			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000024	.000132***	.000175***	.000352***	4.3358	-.2257	-.45589	-7.157**
	(2.3e-05)	(2.8e-05)	(4.6e-05)	(1.1e-04)	(7.45)	(3.31)	(3.35)	(3.53)
Length	.000038*	-.000033	-.000111***	-.000256***	-.068268	-.13934	-.25187	-.57534***
	(2.1e-05)	(2.2e-05)	(3.8e-05)	(9.9e-05)	(.125)	(.186)	(.232)	(.186)
Inv. Odds	.94345***	.93274***	.91721***	.88986***	-48.113	-23.404	-22.343	20.421
	(1.4e-03)	(2.1e-03)	(4.3e-03)	(.016)	(55.2)	(33)	(28.7)	(24.6)
Internet	-.002719***	.000859	-.000235	.000929	6.9006	13.598***	5.6707*	-.036778
	(7.0e-04)	(6.8e-04)	(8.0e-04)	(1.0e-03)	(7.11)	(3.87)	(3.2)	(2.7)
Phone	-.002194	-.003556*	-.000823	.010847	57.835	-1.4502	22.942	12.804
	(2.5e-03)	(2.1e-03)	(3.2e-03)	(9.9e-03)	(75)	(16.3)	(18)	(17.8)
Constant	.007552***	-.000533	.000358	.006768**	17.403	-24.572**	-15.02	2.8449
	(2.4e-03)	(2.3e-03)	(3.5e-03)	(3.4e-03)	(19.8)	(10.5)	(15.2)	(9.03)
N	7,360,949	8,329,414	8,104,150	4,749,204	7,360,949	8,329,414	8,104,150	4,749,204

OLS results. Sample broken by the number of days customer was active. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

does not seem to be any statistically significant effect of stake size on the net revenue, except for the group of the most active bettors, where the effect is statistically significant and negative. Arguably, this can be seen as evidence of betting activity being an inappropriate measure of bettors' skill. In fact, it is not clear why the most active bettors should be labeled as the most skilled, as the image of a professional bettor is of one who selects profitable opportunities and not of one who places many bets every day.[10] Hence, the indirect measure of skills via betting activity does not seem to be a particularly reliable one.

[10] In fact, a bettor who places many bets every day could be viewed as an addicted gambler rather than a professional. From this point of view, the negative effect in column (8) of both regressions would make logical sense.

6.3 Selection into Betting

Because we do not observe the actions of bettors outside of the specific betting company that provided the dataset for this study, there is a possibility that bettors may select into or out of the market based on the circumstances on the market or due to their recent betting experience. In order to examine the possibility of such selection, we repeat the analysis on subsamples of clients broken down by the number of separate years a client enters the sample and based on the longest time interval we observe between two subsequent tickets of each client.

First, we reestimate our baseline specification based on the number of distinct calendar years that the bettor has placed at least one bet. Results of this analysis are presented in Table 6.4. As can be seen from Panel A reporting success regressions, the results are mainly driven by clients who have placed a bet in at least four of the calendar years. This supports the idea that the results are mainly driven by regular bettors, but not necessarily exclusive to betting experts or professionals.

The results on net revenue are included in Panel B of Table 6.4. These results reveal no systematic pattern of stake size on net revenue based on the time the specific client remained in the sample.

Second, we reexamine the baseline estimation eliminating those bettors that had the longest gap between any two tickets longer than some specific period of time. In order to capture several possible time windows, we examine the time frames of 7, 30, 60, 90, 180, and 365 days. Note that as the regression with a time window of 30 days excludes all clients that had the longest gap between placing any two tickets longer than 30 days, it can be labeled as the regression on the sample of clients who

Table 6.4: Regression by Number of Years in the Sample

Panel A: Success	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	4.7e-06	-.000021	.000048	.000147**	.000117***	.000188***	.000139*	.000167***
	(5.0e-05)	(2.9e-05)	(3.6e-05)	(5.8e-05)	(2.8e-05)	(4.5e-05)	(7.4e-05)	(5.1e-05)
Length	.000087	.000034	.00003	-.000092	.00001	-.00007*	-.000171**	-.000185***
	(5.8e-05)	(2.9e-05)	(4.1e-05)	(6.2e-05)	(3.7e-05)	(4.1e-05)	(8.3e-05)	(5.6e-05)
Inv. Odds	.95266***	.94166***	.94124***	.92782***	.93167***	.9266***	.90969***	.90685***
	(3.5e-03)	(2.0e-03)	(3.3e-03)	(7.8e-03)	(3.4e-03)	(3.6e-03)	(8.8e-03)	(7.9e-03)
Internet	-.007368***	-.002554*	-.001575*	.001546	-.001288	-.000797	.001867	.000823
	(2.4e-03)	(1.4e-03)	(8.6e-04)	(1.9e-03)	(1.2e-03)	(1.2e-03)	(1.7e-03)	(6.6e-04)
Phone	-.013096	-.011859**	.001028	-.004087	.000046	-.006505**	.020818**	-.002006
	(.013)	(4.8e-03)	(3.3e-03)	(4.3e-03)	(3.3e-03)	(2.7e-03)	(9.1e-03)	(4.9e-03)
Constant	-.080297***	.015301***	.010027**	.004673	-.004633	-.004035	.005145	.004396
	(6.7e-03)	(5.2e-03)	(4.0e-03)	(4.6e-03)	(4.1e-03)	(3.9e-03)	(3.9e-03)	(3.1e-03)
N	1,133,027	4,127,226	4,037,600	2,007,122	2,444,444	2,907,519	2,783,636	9,103,143

Panel B: Net Revenue	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	47.563*	-19.665***	.31605	-19.745*	18.123**	5.5168	.65584	.26412
	(25.4)	(4.74)	(6.78)	(11.5)	(8.49)	(4.22)	(10.4)	(3.9)
Length	-.18544	-.23969*	-.25187	.16832	-.10827	-.39147	-.19066	-.33444
	(.277)	(.134)	(.163)	(.548)	(.241)	(.333)	(.399)	(.206)
Inv. Odds	-258.74*	112.73***	-20.064	173.85	-215.09**	-67.964	-59.675	-39.169
	(136)	(33.6)	(52.1)	(115)	(102)	(43.5)	(113)	(35.5)
Internet	27.171	-4.1059	7.7202	-5.55	15.689	17.2***	15.382	4.5908
	(28.9)	(5.39)	(5.04)	(12.5)	(14)	(5.19)	(9.68)	(2.84)
Phone	-1138**	237.51**	135.12*	-28.634	-66.567	11.214	132.43	-39.492
	(570)	(99)	(74.2)	(49.5)	(56.7)	(16.6)	(154)	(24.1)
Constant	-83.502***	-7.6816	26.613	7.8272	-13.48	-51.955	-.56462	.12164
	(21.4)	(59.6)	(33.5)	(21.7)	(17)	(39.7)	(16.3)	(6.14)
N	1,133,027	4,127,226	4,037,600	2,007,122	2,444,444	2,907,519	2,783,636	9,103,143

OLS results. Sample broken by the number of calendar years in which the customer has placed a bet. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

Table 6.5: Regression by Maximum Gap Allowed to Stay in the Sample

Panel A: Success Max. Gap in Days	7	30	60	90	180	365
	(1)	(2)	(3)	(4)	(5)	(6)
Stake/100	.000151**	.000144***	.000131***	.000122***	.000101***	.000105***
	(6.5e-05)	(4.5e-05)	(3.1e-05)	(2.5e-05)	(2.3e-05)	(1.9e-05)
Length	.000022	-.0001*	-.000078**	-.000075**	-.000079***	-.000071***
	(6.6e-05)	(5.7e-05)	(3.9e-05)	(3.3e-05)	(2.7e-05)	(2.4e-05)
Inv. Odds	.9313***	.91814***	.92065***	.9231***	.92479***	.92645***
	(4.8e-03)	(7.6e-03)	(4.9e-03)	(4.0e-03)	(3.1e-03)	(2.7e-03)
Internet	-.00008	.000438	-.00023	-.000348	-.000296	-.000268
	(2.4e-03)	(7.7e-04)	(5.9e-04)	(5.2e-04)	(4.6e-04)	(4.2e-04)
Phone	.012741	-.006646*	-.007178***	-.004299*	-.002205	-.002254
	(.014)	(3.6e-03)	(2.6e-03)	(2.3e-03)	(1.9e-03)	(1.6e-03)
Constant	-.00452	.004661*	.003551*	.004277**	.001597	.002213
	(6.2e-03)	(2.8e-03)	(2.1e-03)	(1.9e-03)	(2.0e-03)	(1.7e-03)
N	1,438,609	8,390,353	13,106,809	16,273,535	21,544,410	25,504,980

Panel B: Net Revenue Max. Gap in Days	7	30	60	90	180	365
	(1)	(2)	(3)	(4)	(5)	(6)
Stake/100	60.039**	14.797	5.96	5.9255	-.13802	2.563
	(27.3)	(13.3)	(6.76)	(5.34)	(4.79)	(4.62)
Length	.074264	-.2005	-.11023	-.16203	-.26187**	-.20371*
	(.439)	(.236)	(.183)	(.148)	(.117)	(.115)
Inv. Odds	-395.95**	-128.34	-69.148	-67.515	-20.16	-44.41
	(154)	(90.6)	(52)	(41.5)	(37.4)	(37.5)
Internet	43.282**	14.203*	10.641**	8.3904**	5.4271*	8.1961**
	(19.3)	(8.09)	(5.08)	(3.76)	(3.24)	(3.32)
Phone	-357.14	-63.07	-51.42	-19.093	31.391	31.042
	(456)	(43.2)	(33.1)	(24.5)	(27.4)	(24.4)
Constant	-44.726	-19.857*	-14.794*	-20.104*	-9.7259	-9.9187
	(32.4)	(12)	(8.19)	(11.1)	(11)	(9.21)
N	1,438,609	8,390,353	13,106,809	16,273,535	21,544,410	25,504,980

OLS results. Sample broken by the maximum time interval allowed between two subsequent tickets for the customer to stay in the sample. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

placed a bet at least every month.

The results are presented in Table 6.7. From the success regressions in Panel A, we can see that as the main relationship is valid for all examined time exclusions, selection out of the betting based on bad luck is likely not the driving factor behind our results.

Finally, the net revenue regressions presented in Panel B of Table 6.7 reveal an interesting finding: only those bettors who without exception place at least one bet every single week can consistently vary their stake size in order to increase the net revenue of their bets. This complements our findings that only the bettors with the highest skill level are able to produce the same result.

6.4 Impact of Unusual Wins

7 Conclusion

Our study analyzes the role of stake size in the sports betting market. We utilize a unique sports betting dataset provided by one of the largest betting companies in the Czech Republic. The main advantage of this dataset is access to the betting histories of the company's registered customers.

We find that higher stakes are associated with increased success rates. Thus, individual bettors are able to choose more probable betting events for bets with relatively high stakes. Nevertheless, due to the overall low success rate, the effects of stake size on the net revenue of the bettor are statistically insignificant. Thus, while a typical bettor wins more often when placing a higher stake, he does not on average lose less money. To the best of our knowledge, this study is the first to identify this correlation.

Table 6.6: Regressions - Tickets after Unusual Win - by Multiplication of Average Bet

	Panel A: Success			Panel B: Net Revenue				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stake/100	.000339***	.000326***	.000134***	.000135***	-.34283	-.35879	-3.5883	-3.5889
	(3.0e-05)	(2.9e-05)	(2.0e-05)	(1.9e-05)	(3.08)	(3.12)	(3.82)	(3.82)
$(\text{Stake}/100)^2$	-1.3e-07***	-1.2e-07***	-4.4e-08***	-4.3e-08***	.004618	.004624	.005833	.005833
	(2.5e-08)	(2.4e-08)	(1.4e-08)	(1.4e-08)	(6.1e-03)	(6.2e-03)	(6.5e-03)	(6.5e-03)
Length	-.000052***	-.00005***	-.000059***	-.000071***	-.3826***	-.37552***	-.23551**	-.24334**
	(1.8e-05)	(1.8e-05)	(2.2e-05)	(2.2e-05)	(.083)	(.08)	(.099)	(.1)
Inv. Odds	.93101***	.93104***	.92726***	.9271***	-6.7529	-10.73	3.0129	3.1291
	(2.1e-03)	(2.2e-03)	(2.4e-03)	(2.4e-03)	(21.3)	(21.4)	(30.1)	(30.1)
Internet		-.000535**	-.002604***	-.000189		3.942**	-.35535	4.6227*
		(2.5e-04)	(3.7e-04)	(3.9e-04)		(1.9)	(2.68)	(2.72)
Phone		.003171***	-.0027*	-.002352		9.3195	32.918*	33.774*
		(8.6e-04)	(1.4e-03)	(1.4e-03)		(15.2)	(19.6)	(19.7)
Constant	-.007471***	-.007411***	-.005827***	.002146	-15.336***	-16.253***	-13.586***	-4.8221
	(2.8e-04)	(2.8e-04)	(3.9e-04)	(1.5e-03)	(1.81)	(1.98)	(2.61)	(7.61)
User FE			Yes	Yes			Yes	Yes
Week FE				Yes				Yes
N	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717	28,543,717

OLS results. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

Although the effect of a stake on the success rate is generally valid for all bettors except for those with the lowest skill level, only the bettors with the highest skills can utilize their skill such that they significantly increase their net revenue and thus lose less money.

The decomposition of the bets based on the channel through which they were placed reveals that the effects on the success rate disappear when bets are placed online on the website of the betting company. This may be explained by the fact that because bets placed over the Internet have a generally higher success rate, the effect is less pronounced in their case.

The results are valid regardless of whether individual bettors' fixed effects are included in the analysis, indicating that the relationship between stakes and the

Table 6.7: Regressions - Tickets after Unusual Win at least 20 times higher than Average Bet by Time after the Win

Panel A: Success Time after the Win	7	30	60	90	180	365
	(1)	(2)	(3)	(4)	(5)	(6)
Stake/100	.000151**	.000144***	.000131***	.000122***	.000101***	.000105***
	(6.5e-05)	(4.5e-05)	(3.1e-05)	(2.5e-05)	(2.3e-05)	(1.9e-05)
Length	.000022	-.0001*	-.000078**	-.000075**	-.000079***	-.000071***
	(6.6e-05)	(5.7e-05)	(3.9e-05)	(3.3e-05)	(2.7e-05)	(2.4e-05)
Inv. Odds	.9313***	.91814***	.92065***	.9231***	.92479***	.92645***
	(4.8e-03)	(7.6e-03)	(4.9e-03)	(4.0e-03)	(3.1e-03)	(2.7e-03)
Internet	-.00008	.000438	-.00023	-.000348	-.000296	-.000268
	(2.4e-03)	(7.7e-04)	(5.9e-04)	(5.2e-04)	(4.6e-04)	(4.2e-04)
Phone	.012741	-.006646*	-.007178***	-.004299*	-.002205	-.002254
	(.014)	(3.6e-03)	(2.6e-03)	(2.3e-03)	(1.9e-03)	(1.6e-03)
Constant	-.00452	.004661*	.003551*	.004277**	.001597	.002213
	(6.2e-03)	(2.8e-03)	(2.1e-03)	(1.9e-03)	(2.0e-03)	(1.7e-03)
N	1,438,609	8,390,353	13,106,809	16,273,535	21,544,410	25,504,980
Panel B: Net Revenue Time after the Win	7	30	60	90	180	365
	(1)	(2)	(3)	(4)	(5)	(6)
Stake/100	60.039**	14.797	5.96	5.9255	-.13802	2.563
	(27.3)	(13.3)	(6.76)	(5.34)	(4.79)	(4.62)
Length	.074264	-.2005	-.11023	-.16203	-.26187**	-.20371*
	(.439)	(.236)	(.183)	(.148)	(.117)	(.115)
Inv. Odds	-395.95**	-128.34	-69.148	-67.515	-20.16	-44.41
	(154)	(90.6)	(52)	(41.5)	(37.4)	(37.5)
Internet	43.282**	14.203*	10.641**	8.3904**	5.4271*	8.1961**
	(19.3)	(8.09)	(5.08)	(3.76)	(3.24)	(3.32)
Phone	-357.14	-63.07	-51.42	-19.093	31.391	31.042
	(456)	(43.2)	(33.1)	(24.5)	(27.4)	(24.4)
Constant	-44.726	-19.857*	-14.794*	-20.104*	-9.7259	-9.9187
	(32.4)	(12)	(8.19)	(11.1)	(11)	(9.21)
N	1,438,609	8,390,353	13,106,809	16,273,535	21,544,410	25,504,980

OLS results. Sample broken by the maximum time interval allowed between two subsequent tickets for the customer to stay in the sample. Panel A dependent variable: Whether the ticket won. Panel B dependent variable: Net revenue of the ticket. Fixed effects for individual bettors and weeks are included. Standard errors adjusted for clusters on customer level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Estimation of the model.

probability of winning is driven mostly by the variation within individual bets rather than the variation between the bettors.

Note that one potential caveat of our study is that it takes the ticket lengths and odds composition as given, while these are in fact determined by the decisions of the bettors. Future research is needed to examine these decision-making processes.

References

Andrikogiannopoulou, A. & Papakonstantinou, F. (2011). Market Efficiency and Behavioral Biases in the Sports Betting Market. Working paper, University of Geneva and Swiss Finance Institute, and Imperial College London - Imperial College Business School.

Barberis, N. & Thaler, R. (2003). A Survey of Behavioral Finance. In G. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the Economics of Finance* (1 ed.)., volume 1, Part 2 chapter 18, (pp. 1053–1128). Elsevier.

Braun, S. & Kvasnicka, M. (2013). National Sentiment and Economic Behavior: Evidence From Online Betting on European Football. *Journal of Sports Economics*, *14*(1), 45–64.

Feess, E., Müller, H., & Schumacher, C. (2014). The favorite-longshot bias and the impact of experience. *Business Research*, *7*(2), 217–234.

Feess, E., Müller, H., & Schumacher, C. (2016). Estimating risk preferences of bettors with different bet sizes. *European Journal of Operational Research*, *249*(3), 1102 – 1112.

Franklin, A., Morris, S., & Shin, H. S. (2006). Beauty Contests and Iterated Expectations in Asset Markets. *The Review of Financial Studies*, *19*(3), 719–752.

Gainsbury, S. M. & Russell, A. (2015). Betting patterns for sports and races: A longitudinal analysis of online wagering in australia. *Journal of Gambling Studies*, *31*(1), 17–32.

Golec, J. & Tamarkin, M. (1998). Bettors Love Skewness, Not Risk, at the Horse Track. *Journal of Political Economy*, *106*(1), 205.

Grant, A. (2013). Betting on Simultaneous Events and Accumulator Gambles. In D. S. Siegel & L. Vaughan-Williams (Eds.), *The Oxford Handbook of the Economics of Gambling* (pp. 341–369). Oxford University Press.

Kopriva, F. (2015). Constant Bet Size? Don't Bet on It! Testing Expected Utility Theory on Betfair Data. CERGE-EI Working Papers wp545, The Center for Economic Research and Graduate Education - Economics Institute, Prague.

Matarazzo, O., Carpentieri, M., Greco, C., & Pizzini, B. (2018). Are the Gambler's Fallacy or the Hot-Hand Fallacy due to an Erroneous Probability Estimate? In *Multidisciplinary Approaches to Neural Computing* (pp. 353–368). Springer.

Moskowitz, T. J. (2015). Asset Pricing and Sports Betting. Chicago Booth Research Paper No. 15-26.

Ottaviani, M. & Sørensen, P. N. (2008). The Favorite-Longshot Bias: An Overview of the Main Explanations. In D. B. Hausch & W. T. Ziemba (Eds.), *Handbook of Sports and Lottery Markets* (pp. 83ďż″–101).

Sauer, R. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, *36*(4), 2021–2064.

Snowberg, E. & Wolfers, J. (2010). Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy*, *118*(4), 723–746.

Staněk, R. (2017). Home bias in sport betting: Evidence from Czech betting market. *Judgment and Decision Making*, *12*.

Zafiris, N. (2014). When Is a Multiple Bet Better than a Single? *Journal of Gambling Business & Economics*, *8*(2), 1 – 15.