CERGE
Center for Economics Research and Graduate Education
Charles University

Essays on Sports Economics

Radek Janhuba

Dissertation

Prague, September 4, 2018

Radek Janhuba

Essays on Sports Economics

Dissertation

Prague, September 4, 2018

Dissertation Committee

Jan Hanousek (CERGE-EI; chair)

Randall K. Filer (CUNY Hunter College)

Stepan Jurajda (CERGE-EI)

Nikolas Mittag (CERGE-EI)

Referees

Brad R. Humphreys (West Virginia University)

Daniel Rees (University of Colorado Denver)

Contents

Abstract v

Abstrakt vii

Acknowledgments ix

Preface 1

1 Do Victories and Losses Matter? Effects of Football on Life Satisfaction 5
- 1.1 Introduction . . . 6
- 1.2 Literature Review . . . 9
  - 1.2.1 Psychological Effects of Sports . . . 10
  - 1.2.2 Sports and Subjective Well-Being . . . 11
- 1.3 Data . . . 13
  - 1.3.1 Football Results . . . 13
  - 1.3.2 Behavioral Risk Factor Surveillance System . . . 16
  - 1.3.3 Linking Games to Survey Responses . . . 16
- 1.4 Methodology . . . 18
- 1.5 Results . . . 21
  - 1.5.1 Baseline Analysis . . . 21
  - 1.5.2 Social Context: Sharing the Wins Together . . . 25
  - 1.5.3 Demographic Effect Heterogeneity . . . 27
- 1.6 Robustness and Sensitivity Checks . . . 30
  - 1.6.1 Selection of Cutoff Values . . . 30
  - 1.6.2 Specifics of Empirical Methodology . . . 32
- 1.7 Conclusion . . . 36

Appendices 39
- 1.A Appendix 1.A: Composition of Control Variables . . . 39
- 1.B Appendix 1.B: Teams Used in the Analysis . . . 42
- 1.C Appendix 1.C: Tables with All Football Covariates . . . 44
- 1.D Appendix 1.D: Full Regression Results . . . 49

2 **Criminals on the Field: A Study of College Football** 57
2.1 Introduction 58
2.2 Football Specifics and Intervention Details 61
2.2.1 The Game of Football 61
2.2.2 Change in the Number of Officials 62
2.3 Data 65
2.4 Methodology 67
2.5 Results 70
2.5.1 All Penalties 70
2.5.2 Areas of Officiating Coverage 71
2.5.3 Experience with the Policy 73
2.5.4 Role of Team Quality 76
2.5.5 Robustness Checks 78
2.6 Conclusion 81

Appendices 83
2.A Appendix 2.A: Number of Top Teams 83
2.B Appendix 2.B: Logit Results 87

3 **High Bets for the Win? The Role of Stake Size in Sports Betting** 93
3.1 Introduction 94
3.2 The Betting Market 97
3.2.1 Single and Accumulator Bets 98
3.2.2 Literature Review 100
3.3 Data 101
3.4 Methodology 104
3.5 Results 105
3.5.1 Baseline Analysis 105
3.5.2 Role of Skills 107
3.5.3 Channels 108
3.5.4 Relationship with Ticket Length and Odds 111
3.6 Robustness Checks 113
3.6.1 Non-Linear Relationship 113
3.6.2 Alternative Measures of Skill 114
3.6.3 Selection into Betting 115
3.7 Conclusion 119

Abstract

In the first chapter, I examine the effects of emotional shocks on subjective well-being and the role social context plays in how shocks are experienced. Using data from the Behavioral Risk Factor Surveillance System (BRFSS), the study uses an ordered logit model to estimate the effects of the local college football team's wins and losses on the life satisfaction of local citizens. The analysis suggests that unexpected wins have positive effects on life satisfaction. The results are driven entirely by games played at the home stadium, indicating that the impacts of emotional shocks are larger if the experience is shared with other fans. Moreover, the effects increase with the size of the stadium relative to the local population, suggesting that social context is likely to be the underlying factor. Surprisingly, no effects are found for cases of unexpected losses.

The second chapter examines the relationship between the number of on-field officials and committed fouls, a phenomenon connected to the economics of crime. Economists have found mixed evidence on what happens when the number of police increases. On one hand, more law enforcers means a higher probability of detecting a crime, which is known as the monitoring effect. On the other hand, criminals incorporate the increase into their decision-making process and thus may commit fewer crimes, constituting the deterrence effect. This study analyzes the effects of an increase in the number of on-field college football officials, taking players as potential

criminals and officials as law enforcers. Analyzing a novel play-by-play dataset from two seasons of college football, we report evidence of the monitoring effect being present in the overall dataset. This effect is mainly driven by offensive penalties that are called in the area of jurisdiction of the added official. Decomposition of the effect indicates the presence of the deterrence effect in cases of penalties with severe punishment or those committed by teams with moderate to high ability, suggesting that teams are able to strategically adapt their behavior following the addition of an official.

In the third chapter, 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 the stake and betting outcomes is driven by variation in individual bets.

Abstrakt

První kapitola zkoumá roli emocionálních šoků na subjektivní ohodnocení blahobytu občanů a to, jakou roli v prožívání těchto šoků hraje společenský kontext. Studie využívá data z dotazníkového šetření Behavioral Risk Factor Surveillance System (BRFSS) a pomocí pořádkového logitu odhaduje efekty výsledků lokálního fotbalového týmu na spokojenost se životem místních obyvatel. Výsledky ukazují, že neočekávané výhry mají na spokojenost se životem pozitivní efekt. Výsledky jsou plně hnány zápasy hranými na domácím stadionu, což ukazuje, že efekty emocionálních šoků jsou silnější, pokud jsou prožívány společně s ostatními fanoušky. Toto zjištění je podpořeno tím, že je efekt rostoucí v relativní velikosti stadionu oproti počtu místních obyvatel. Překvapivým výsledkem je zjištění, že neexistuje žádný efekt neočekávaných proher.

Ve druhé kapitole zkoumáme vztah mezi počtem rozhodčích a faulů, což je vztah dotýkající se ekonomie kriminality. Ekonomové doposud našli nejednoznačné výsledky při zkoumání vlivu počtu policistů na spáchané přestupky. Na jednu stranu zvýšená koncentrace policistů zvyšuje pravděpodobnost odhalení porušení zákona, což je označováno jako monitorovací efekt. Na druhou stranu potenciální zločinci tento nárůst zohlední do svého rozhodování a můžou tak ve výsledku páchat méně přestupků, což je nazýváno jako odrazující efekt. Tato studie analyzuje efekty navýšení počtu rozhodčích na hřišti v zápasech amerického fotbalu, přičemž rozhodčí jsou

pro její účely bráni jako policisté a hráči jako potenciální zločinci. Studie analyzuje nově zkonstruovaný datový soubor pokrývající dvě sezóny univerzitního fotbalu a na celém vzorku nachází přítomnost monitorovacího efektu. Výsledky jsou hnány zejména fauly v oblasti, která je sledována nově přidaným rozhodčím. Dekompozice efektů poukazuje také na přítomnost odrazujícího efektu, a to v případě závažných faulů a v případě faulů spáchaných relativně výkonnostně silnými týmy. Tyto výsledky naznačují, že jsou týmy po přidání rozhodčího schopny strategicky měnit své chování.

Ve třetí kapitole analyzujeme vliv velikosti vsazené částky ve sportovním sázení. Naše hlavní výzkumná otázka je, zda výše vsazené částky predikuje výsledky sázení. Konkrétně ve studii zkoumáme, zda jsou sázkaři schopni konzistentně vybírat výnosnější sázky v nejdůležitějších momentech. Studie využívá unikátní data obsahující více než 28 miliónů reálných sázek vsazených registrovanými klienty v české sázkové kanceláři. Výsledky ukazují, že sázkaři jsou schopni měnit vsazenou částku tak, aby zvýšili pravděpodobnost výhry, avšak nikoliv až tak, aby zároveň zvýšili svou čistou pozici vyplývající ze sázkové aktivity. Výsledky nadále naznačují, že zlepšení své čisté pozice jsou schopni pouze nejschopnější sázkaři. Výsledky studie jsou platné bez ohledu na to, zda do analýzy zahrneme fixní efekty jednotlivých sázkařů, což ukazuje, že vztah mezi velikostí vkladu a výsledky sázení je hnaný variací jednotlivých sázek.

Acknowledgments

This dissertation was completed with the help of numerous people to whom I wish to express my gratitude.

First, I thank my supervisor, Jan Hanousek, for valuable guidance throughout my PhD studies. Jan was extremely supportive in all phases of my research projects and has provided excellent advice on all of the studies contained here. I also had the opportunity to work for him as a research assistant, which led me to acquire essential skills for data management needed to execute the second study in this dissertation.

Other members of the dissertation committee also deserve my gratitude for the tremendous support and valuable time they spent advising me. Stepan Jurajda provided helpful consultations on all three studies in this dissertation. We also coauthored a study, the experience of which increased the quality of the papers contained in this dissertation. Randall Filer and Nikolas Mittag also provided valuable feedback on all of the studies contained here.

Two chapters in this dissertation are coauthored, and I am grateful to my fellow writers for their time and effort spent on these two studies. This dissertation would not be complete without the work by Kristyna Cechova and Jakub Mikulka.

During my studies, I had an opportunity to spend a semester as a visiting scholar at West Virginia University. This stay would not have been possible without the support of Brad Humphreys, who arranged the necessary details. The discussions, consultations, and seminars that I engaged in at WVU were beneficial in providing insights for the studies in this dissertation.

In addition to colleagues and peers already mentioned, the studies in this manuscript benefit from discussions with numerous people. Specifically, I thank William Appleman, Michal Bauer, Libor Dusek, Kamil Kovar, Stefan Lyocsa, Bryan McCannon, Neil Metz, Phil Miller, Jakub Steiner, Danko Tarabar, Nicholas Watanabe, Pamela Wicker, Christopher Yencha, Jan Zapal, and several anonymous referees for helpful comments and suggestions. Participants at numerous conferences, seminars, and workshops also provided valuable feedback.

This dissertation would not be as polished without the tremendous help of the Academic Skills Center at CERGE-EI, which provided academic writing training and performed the English editing of this thesis. In particular, I thank Andrea Downing, Grayson Krueger, and Deborah Novakova.

The study in the first chapter was supported by GAUK project No. 162415, and the whole dissertation was completed with institutional support RVO 67985998 from the Czech Academy of Sciences. My stay at West Virginia University received support from the Nadani foundation. All financial support is hereby gratefully acknowledged.

All errors remaining in the text are my own.

Prague, Czech Republic | Radek Janhuba

September 6, 2018

Preface

This dissertation contains three essays on sports economics, a rapidly growing field of economics. Sports generally have a useful property of being relatively well measured, and, particularly after the recent advances in automatic data collection, information on sporting outcomes has become accessible. Moreover, the existence of sports betting markets assures that not only information about the outcomes, but also about the ex ante expectations of these outcomes is available. The combination of these useful properties has recently resulted in data from sports events and competitions becoming an increasingly common kind of data analyzed in empirical studies.

The three studies contained in this dissertation provide three applications of data from the domain of sports on economic research questions. These three studies contribute to the economics of well-being, crime, and betting. In what follows of the Preface, the three chapters are briefly described. Note that in order to distinguish this introductory and motivative text from the Abstract, I intentionally abstain from discussing the results of the specific studies here in the Preface.

The first chapter examines the role of social context in how emotional shocks are experienced. Specifically, it examines the impact of unexpected results of the local college football team on the subjective well-being measure of life satisfaction, which is represented by survey responses in areas where the particular team has substantial fan support. Previous studies examining the relationship between the subjective well-being and football results suffered from the inability to sufficiently identify which respondents follow which team. This study provides a novel way of combining the survey responses with teams using Facebook *likes*. By knowing the county in which the respondent lives and utilizing the percentage of all Facebook *likes* of top teams in each zip code area collected by *The New York Times*, I can identify counties where the majority of fans supports one specific team. I then use the timing of the interview to link each survey response to the previous game of the particular team. The empirical analysis then concentrates on the effects of unexpected results, conditioning on the pre-game betting market's expectations about the outcome. In order to explore the role of social context in experiencing emotional shocks, I distinguish whether the game in question was played at the home stadium or elsewhere.

The second chapter is a policy evaluation and analyzes an intervention in which the National Collegiate Athletic Association (NCAA) added an on-field official in the highest college football division, thus increasing the number of officials from seven to eight. The setting of the study takes officials as law enforcers (police) and players as potential criminals. When the number of police increases, the police observe crime better and the probability of catching a lawbreaker increases, which is known as the monitoring effect. However, potential criminals incorporate this increase into their decision-making process and may consequently commit fewer crimes, constituting the deterrence effect. Our study contributes to the ongoing discussion on the existence and strength of these two effects. Like the study in the first chapter, the research combines several separate data sources to analyze the research question in a play-by-play setting. We exploit information on the specific crew of officials as well as the skills of teams playing in the particular game. The study is the first to analyze the NCAA intervention on a nation-wide dataset and to concentrate on the time period during which the policy change was implemented universally.

The third chapter examines the behavior of bettors on the betting market. Specifically, we focus on the effect of stake size on betting outcomes, by which we take the probability of a specific bet winning and its net revenue. The study uses a unique dataset containing bets that were actually placed at a bookmaking company in the Czech Republic, thus allowing us to analyze actual transaction-level data rather than only price information. We utilize the decisions of bettors to combine individual bets into accumulator (parlay) tickets. For such a bet to win a positive amount, all of the individual opportunities have to win. We show that even though accumulator bets carry a lower expected return and higher variance due to the margin of the betting company, they are extremely popular. We exploit the fact that almost all clients regularly place accumulator bets, and we use the number of opportunities on a betting ticket as a control variable in the analysis. This study is the first to employ information from accumulator bets in the context of actually placed bets, and is also the first to empirically examine the role of stake size in the betting market.

Chapter 1

Do Victories and Losses Matter? Effects of Football on Life Satisfaction

Radek Janhuba[1]

Abstract

This study examines the effects of emotional shocks on subjective well-being and the role social context plays in how shocks are experienced. Using data from the Behavioral Risk Factor Surveillance System (BRFSS), this paper uses an ordered logit model to estimate the effects of a local college football team's wins and losses on the life satisfaction of local citizens. The analysis suggests that unexpected wins have positive effects on life satisfaction. The results are driven entirely by games played at the home stadium, indicating that the impacts of emotional shocks are larger if the experience is shared with other fans. Moreover, the effects increase with the size of the stadium relative to the local population, suggesting that social

[1]An earlier version of this work was published in Janhuba, R. (2016) "Do Victories and Losses Matter? Effects of Football on Life Satisfaction", *CERGE-EI Working Paper Series No. 579*. The study was supported by Charles University, GAUK project No. 162415, and with institutional support RVO 67985998 from the Czech Academy of Sciences. I thank Michal Bauer, Randall Filer, Jan Hanousek, Brad Humphreys, Stepan Jurajda, Neil Metz, Nikolas Mittag, Nicholas Watanabe, participants in the 2015 MVEA Kansas City, 2016 YEM Brno, and 2017 WEAI conferences, and participants in seminars at CERGE-EI, WVU, Syracuse, and Technical University of Ostrava for helpful comments and suggestions. All remaining errors are my own.

context is likely to be the underlying factor. Surprisingly, no effects are found for cases of unexpected losses.

1.1 Introduction

This study examines the effects of emotional shocks on subjective well-being (henceforth SWB). Specifically, we examine the effects of a local college football[2] team's wins and losses on responses to the life satisfaction question, measuring the overall, long-term level of satisfaction with one's life. We are particularly interested in the role social context plays in experiencing these emotional shocks. Hence, we link the literature on the effects of emotional shocks caused by sports (Card and Dahl, 2011; Eren and Mocan, 2018) with studies examining the behavioral effects of group identity (Charness et al., 2007; Depetris-Chauvin et al., 2018).

We examine whether the effects of football results on SWB are magnified when the experience is shared with other fans. We exploit the fact that football games are played at home as well as on the road. While fans usually watch road games on TV, many attend home games in person. Moreover, during the home-game days, the stadium surroundings are impacted by an influx of fans, tailgate parties, and other phenomena associated with the event. Thus, even for people who do not attend the game, being present in the stadium surroundings involves one in the social environment around the game.

To better understand the relationship between sports and SWB, it is important to point out that our research interest lies in observing whether *unexpected* outcomes matter.[3] Thus, the methodology of the study is constructed so as to allow us to distinguish between unexpected and *general* outcomes, which we define as results which are not surprising.[4]

The examination of unexpected outcomes is motivated by two economic concepts.

[2]Note that throughout this study, the word *football* indicates specifically American football. When needed, the standard, European football, is referred to as *soccer* (derived from its full name *association football*).

[3]We define unexpected results based on the pre-game betting market valid in Las Vegas at kickoff time. See Section 1.3.1 for more information.

[4]Note that it is not possible to label such outcomes as *expected*, because unexpected results are defined as having been a result that carries a sufficient level of surprise.

First, based on the theory of reference dependent preferences, rational agents are expected to form expectations with respect to information available ex ante (Koszegi and Rabin, 2006). In our setting, this means that fans' emotions are likely to be influenced differently when a result carries an element of surprise (relative to the benchmark formed by the expectations) and when it does not. The incorporation of the unexpectedness of the results into the analysis may thus be viewed as empirical validation of the reference-point utility of Koszegi and Rabin (2006).

Second, millions of Americans attend sports events every week and tens of millions watch sports on TV. While sports events undoubtedly generate a great deal of entertainment value, the suspense and surprise model by Ely et al. (2015) suggests that unexpectedness is the main driving factor behind the entertainment value derived. Thus, when analyzing shocks induced by sports events, it is necessary to incorporate the unexpected component of the results into the empirical methodology.

This study focuses on results from American college football, which has an extremely strong fan base.[5] Previous research has shown that being a sports fan is associated with one's emotions (Kerr et al., 2005; Jones et al., 2012). Hence, college football is likely to have strong ties to the emotional domain. Moreover, individual wins matter in college football. With only 12 regular season games each year and 4 of 130 teams reaching the playoffs, the marginal effect of each individual result is stronger than in all other major sports. Thus, unexpected football outcomes subject fans to a relatively strong emotional shock.

The contributions of this study are threefold. First, to our knowledge, this is the first study to examine the social context of the psychological effects of sports. While previous research has found that group identification influences behavior (Charness et al., 2007), psychological effects of groups have thus far been studied experimentally (see Kugler et al. 2012 for an overview). In this sense, this study provides novel field evidence on the economic psychology of groups.

Second, we implement a novel methodology of using data from Facebook *likes* to match teams with their fans (see Section 1.3.3). We are currently unaware of

[5]Market research for 2012 estimated that 43% of the US population followed college football. Source: http://sportsaffiliates.learfieldsports.com/files/2012/11/College-vs.-Pro.pdf

any other study that has used Facebook *likes* to link two separate datasets in an analogical way, making our approach novel. While previous work examining the effects of sports has concentrated on data from metropolitan areas, this methodology allows us to use data from non-urban areas as well.

Third, while previous studies on sports and SWB have concentrated on one-off, large-scale tournaments, this study seeks to identify the connection on a dataset utilizing regular weekly games. This eliminates the possibility of a spurious one-time effect that may have taken place around tournaments examined in previous studies. To our knowledge, this study is the first to examine such a relationship in the context of sports and SWB.

We find that unexpected wins in home games have systematic effects on the reported life satisfaction of US residents. More importantly, the results indicate that social context plays an important role in SWB evaluation. Specifically, rather than simply being a fan, it is the effect of being a fan and at the same time sharing the experience of an unexpected win with others that influences the life satisfaction responses. This notion is supported by the fact that areas with higher capacity stadiums relative to the local population are associated with stronger effects.

In terms of magnitude, following an unexpected win at the home stadium, the probability of a respondent reporting the highest life satisfaction category grows by approximately 12 percentage points. Further, back-of-the-envelope calculation suggests that the true value of the effect lies between 12 and 27 percentage points. Nevertheless, although the effect is sizable ex-post for several days following an unexpected win, its overall magnitude is negligible. Thus, it does not endanger comparisons of life satisfaction levels across regions and/or time.

The analysis also finds that there are no effects of unexpected losses, a result that is very surprising in terms of knowledge of sports and psychological processes, where unexpected losses but not unexpected wins were found to influence domestic violence (Card and Dahl, 2011) and judicial sentence lengths (Eren and Mocan, 2018). The stark distinction between these and our findings is likely caused by the nature of the outcomes examined.

Specifically, while domestic violence and judicial sentences are connected to actions that fall within the negative side of emotional scale, we examine a variable linked to positive emotional shocks. Thus, while Card and Dahl's (2011) and Eren and Mocan's (2018) data are likely insensitive to small positive changes in individual well-being, our dataset mainly comprises of life satisfaction evaluations on the positive side of the emotional scale and is therefore likely less sensitive to small negative changes in SWB. Hence, our results present complementary rather than substitute evidence to the results of Card and Dahl (2011) and Eren and Mocan (2018). The combined implication of these results is that unexpected football results in both directions may affect decisions in the connected emotional domain but do not alleviate the general benchmark level of these decisions in the absence of unexpected shocks.[6]

In terms of psychological research concerning changes in well-being, our results may also be seen as complementary evidence to the experimental study of Yechiam et al. (2014), who find that in cases of one-shot interactions, people tend to report greater valuations of gains compared to losses. Because a particular football team usually does not experience many instances of unexpected results throughout a season, unexpected wins and losses can be seen as one-shot events.[7]

The remainder of this paper is structured as follows. Section 1.2 contains a brief literature review. Section 1.3 presents the data used in the estimation. Section 1.4 explains the methodology used in the analysis. Section 1.5 shows empirical results and discusses their importance. Section 1.6 discusses the robustness of our results to alternative specifications. Section 1.7 concludes.

1.2 Literature Review

Most of the previous literature on the effects of sports has concentrated on stadiums and arenas and is not reviewed here. The conclusion of this literature is that stadiums where sports are played do not convey immediate economic benefits to the areas where they are built. For a thorough review of these studies, see Coates and

[6]By the connected emotional domain, we mean outcomes generally associated with positive feelings in cases of unexpected wins and vice versa.

[7]Yechiam et al. (2014) also present evidence that reporting feelings about wins and losses is not necessarily associated with behavioral biases. This can explain why our results seemingly go against the loss aversion theory.

Humphreys (2008).

Several studies have analyzed the economic effects of college sports on local economies. Baade et al. (2008) estimate the economic impact of home college football games and find no evidence of measurable effects. In a follow-up study, Baade et al. (2011) do find a positive effect of home football games for the city of Tallahassee, Florida, suggesting that the local economy gains approximately $2 Million following each football game played at the local stadium. This amount is relatively small compared to the amounts in public subsidies college football teams usually receive. Moreover, Baade et al. (2011) provide evidence that part of the increased revenue comes from a substitution effect within the state of Florida, further diminishing the estimated real economic value-added of organizing college football games.

A stream of literature, e.g. Ahlfeldt and Maennig (2010), Ahlfeldt and Kavetsos (2014), has found that property prices in the surroundings of stadiums rise following stadium construction, suggesting the presence of beneficial intangible effects of sport arenas. Humphreys and Nowak (2017) show that property values in the vicinity of Seattle's arena rose after the Seattle Supersonics moved to Oklahoma, indicating that the team had a detrimental effect on the local community. Although this study focuses on a different topic, these results may serve as an indicator of asset prices incorporating intangible benefits created by the presence of sports teams.

1.2.1 Psychological Effects of Sports

A branch of literature explores situations in which sport enters the psychological domain of agents, which in turn translates into their actions having an "unrelated" impact.

Card and Dahl (2011) find that the reported number of domestic assaults rises significantly in the three hours after a professional football game which the local team unexpectedly lost. Rees and Schnepel (2009) obtain similar results in a sample of Division I college football games and extend its validity to a range of other criminal behavior in the town where the game is played.

Eren and Mocan (2018) analyze juvenile court decisions in Louisiana and find that unexpected losses of the LSU football team lead to increased sentence lengths during the week following the game. Moreover, they find that the results are entirely driven by the portion of judges who attended the LSU University. The study of Eren and Mocan (2018) serves as a strong example of football results influencing a seemingly unrelated phenomenon through the affected psychological domain of decision makers.

Several studies have also found effects following wins of the local team. Agarwal et al. (2013) find evidence of mortgage loan approval rates increasing by more than four percentage points following a large sports event leading to positive sentiment in affected counties. Fernquist (2000) finds that local teams making the playoffs lead to a lower suicide rate in the local population. Chen (2016) observes that immigration judges on average grant an additional 1.5% of asylum petitions on Mondays after the city's professional football team won compared to a loss. Healy et al. (2010) show that the probability of incumbents' reelection in the county of a college football team is approximately 1.5% higher if the particular team wins a game in the 10 days prior to the election.

A distinct stream of literature has focused on the effects of sport teams on stock markets. Edmans et al. (2007) find that individual sentiment following a national team's loss in various sports leads to an abnormal negative return on the affected country's stock exchange. Drake et al. (2016) find that investors' distraction during the NCAA basketball tournament (known as the *March Madness*) creates stock disruptions that are present in the market for a period of 30 to 60 days.

1.2.2 Sports and Subjective Well-Being

To our knowledge, few studies have examined the relationship between sports events and life satisfaction. Most of the existing research linking the two has concentrated on the effects of practicing sports on SWB and is not surveyed here.[8]

The earliest study to observe the effects of sports events on life satisfaction is Schwarz et al. (1987), who found that German males reported a higher general life satisfaction after a 1982 World Cup soccer game that ended with a German win. Although their sample size is very limited, with only 55 observations, the authors

[8] See Section 2 of Kavetsos and Szymanski (2010) for an overview.

conclude that this is an example of momentary happiness transcending into the long-term evaluation, implying the existence of the phenomenon this study aims to identify.

Kavetsos and Szymanski (2010) examine data from 12 European countries to observe whether hosting an important tournament or having an unexpectedly successful national soccer team in a significant tournament, such as the Olympic games or the FIFA World Cup, have overall effects on life satisfaction reported by the country's citizens. Although their study finds limited evidence that the success of the national team has positive implications for inhabitants' life satisfaction, they do find a significant positive effect of hosting a large soccer tournament.

Süssmuth et al. (2010) analyze citizens' willingness to pay for the 2016 FIFA World Cup tournament that took place in Germany. Their results reveal that the reported willingness to pay increased ex post as compared to the same respondents' valuation ex ante, and also indicate that almost 85% of German citizens thought that hosting the FIFA World Cup brought overall net benefits to the country (Süssmuth et al., 2010, p. 208). This is consistent with the findings of Allmers and Maennig (2009), who report a rise in international perception of Germany following the 2016 FIFA World Cup.

A recent study by Depetris-Chauvin et al. (2018) links sports to psychological effects based on national identification. Specifically, Depetris-Chauvin et al. (2018) examine the effects of national soccer teams' results on violence in Africa. Examining large-scale survey data, the study finds that individuals interviewed following their national team's victory are more likely to trust people of other ethnicities. Moreover, Depetris-Chauvin et al.'s (2018) results show that teams that closely qualify for the African Cup of Nations are subject to a lower subsequent degree of violence compared to the countries that did not qualify for the tournament.

Doerrenberg and Siegloch (2014) examine whether being interviewed before or after an international soccer tournament has implications on several dependent variables, using a panel of unemployed individuals in Germany. Although the evidence is mixed for the case of life satisfaction, the study finds a significant decrease in general worries about the economic situation as well as a significant increase in the perceived intention to find work again.

Although the studies described above analyze the effects of sports events on life satisfaction, there is a distinction between their and our approach. Namely, the previous work concentrates on short-timed, large scale tournaments, while this study examines the relationship on data from regular, week-to-week games. This eliminates the possibility of a spurious one-time effect that may have taken place around tournaments examined in previous studies. Moreover, the sample size associated with a large scale dataset allows us to examine potential heterogeneity of the effect in various decompositions, such as those based on differences in demographic characteristics or the extent to which the result was surprising (see Section 1.5.3 for more details). To our knowledge, this study is the first to examine the phenomenon in such settings.

1.3 Data

This section first introduces the two sources of data: football results and the BRFSS survey, which includes the dependent and control variables. Section 1.3.3 follows with a description of the novel method linking these two datasets.

1.3.1 Football Results

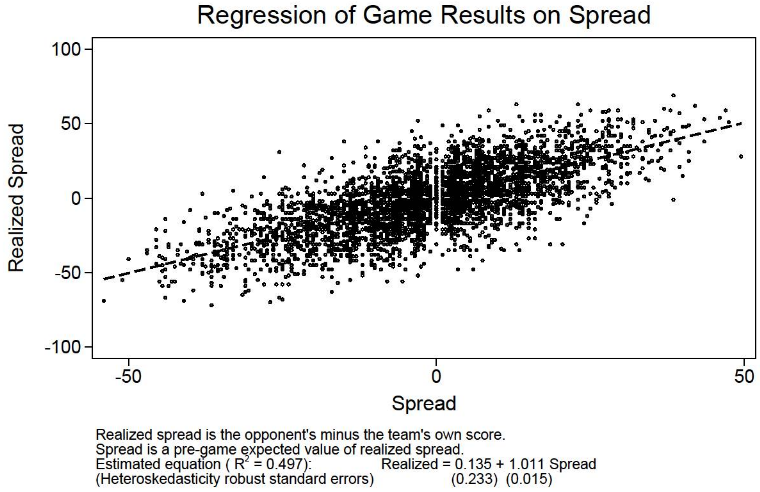
The data on football games were purchased from The Logical Approach[9] and contain betting information available on the Las Vegas market at the kickoff time of each FBS[10] college football game. As the second data set includes surveys conducted from 2005 to 2010, the sample consists of games played between 29th December 2004 and 28th December 2010.

The information about the expected result of a game is included in the *spread*, quoted as the expected number of points to equalize the two opponents valid on the Las Vegas betting market at kickoff time. For example, a spread of -10 means that the team was expected to win the game by 10 points (consequently, the opponent would have the spread quoted as +10 and be expected to lose the game by 10 points).

[9] http://www.thelogicalapproach.com/

[10] FBS is the highest level of college football played in the United States.

Figure 1.3.1: Predicted vs. Realized Spreads



Previous research (e.g. Sauer 1998, Fair and Oster 2007, and Song et al. 2007) has shown that spreads contain the most relevant information that is available ex ante about the outcome of a football game. Our data is consistent with their conclusions, as the regression estimate of the realized value of the spread on its value yields a coefficient of 1.01 with a standard deviation of 0.015, a level that is statistically not significant from the market-efficient value of 1 (see Figure 1.3.1). Therefore, we can use the spread to control for the ex ante probability of a particular team winning the game.

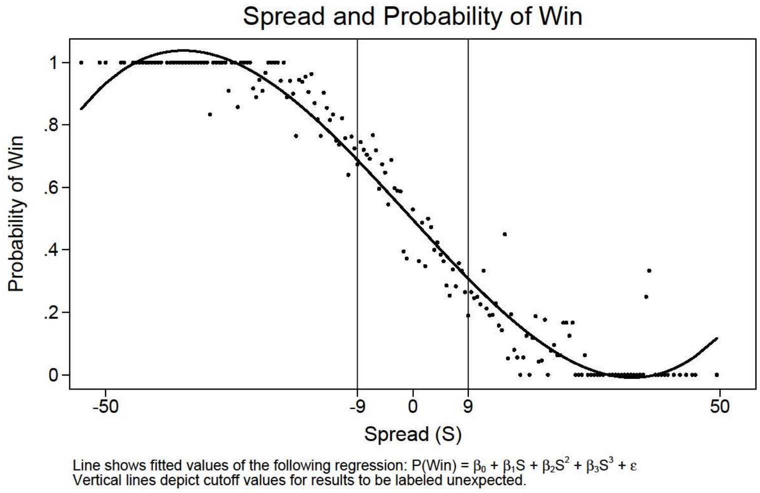
Table 1.3.1: Frequencies of Games by Cutoff Spread

Spread	No.	Col %	Cum %
Lower or equal -9 points	2,182	25.4	25.4
Between -9 and 9 points	4,296	50.1	75.6
Higher or equal 9 points	2,096	24.4	100.0
Total	8,574	100.0	

Source: Author's computation based on games from 2005 until 2010.

A result is defined as *unexpected* if it goes against the spread of 9 points or more in an absolute value. This specific value was selected as it breaks the set of games to approximately one quarter below and above the threshold (see Table 1.3.1), ensuring that the surprise effect is sufficiently strong, while still keeping enough games to allow

Figure 1.3.2: Spread and Probability of Win



for a sizable number of unexpected results. In this sense, the selection is very similar to Card and Dahl's (2011) study, which uses 4 points on NFL data associated with a lower volatility of spreads.[11] In fact, the 75th percentile in their data is equal to 4 points, making our selection comparable after accounting for the difference in the volatility of spreads between the two competitions. Moreover, 9 points is especially useful from the view of football rules, as it is the lowest point difference in a two-possession game.[12] Nevertheless, our empirical results are robust to the selection of this upset threshold (see Section 1.6.1).

Figure 1.3.2 shows the probability a team will win the game based on the spread. The expected probability of winning is less than or equal to 36.4% once the spread is higher than or equal to 9. The probability of an unexpected loss is less than or equal to 39.2% for unexpected losses with a spread lower than or equal to -9.[13]

[11] NFL (National Football League) is the major professional football league in the United States.

[12] In football, when a team scores a touchdown, it receives six points. It then attempts one more play (called "point after try") for which it receives zero, one, or two points. Therefore, once the point difference reaches 9 points, the trailing team has to score at least twice to win the game.

[13]Note that these values present the average probability of a surprise result and do not account for differences in team characteristics. Generally, the probability of an unexpected win will be very low for successful teams that almost never lose, as they will extremely rarely be expected to lose the game by a sufficient margin.

1.3.2 Behavioral Risk Factor Surveillance System

The second data source used is the *Behavioral Risk Factor Surveillance System* (BRFSS), collected daily by the *Centers for Disease Control and Prevention* (CDC) on a wide-ranging sample of American citizens, resulting in a yearly sample size of about 400,000 observations.

The BRFSS is a system of telephone surveys that collects data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive health services. Although the repeated cross-sectional nature of the data inevitably leads to an issue of unobserved heterogeneity, the BRFSS has three main advantages which make it very convenient for our particular setting. First, from 2005 to 2010,[14] the survey contained a life satisfaction question in which respondents self-evaluate themselves on a scale from 1 to 4 by answering the question *"In general, how satisfied are you with your life?"*, with options labeled (from 1 to 4) "very satisfied", "satisfied", "dissatisfied" and "very dissatisfied".[15]

Second, the data set contains FIPS county codes, allowing a much closer geographic link than in the case of data sets which only contain state level identification. As there are multiple FBS football teams in most states, we need such information to match the particular observation to the appropriate team.

Third, the availability of the exact survey date allows us to identify whether the local football team had won or lost the game prior to the survey.

1.3.3 Linking Games to Survey Responses

The crucial question after obtaining the data on survey responses and football games is how to link a specific game to a particular observation (it is straightforward that it may not be sufficient to simply take the closest geographical team to the area where the respondent lives). As mentioned in the introduction, our method uses data from Facebook *likes*. Specifically, it looks at which team has the largest share

[14] Since 2011, the question has been moved into the optional part of the questionnaire and is asked in only a small number of states.

[15] Throughout this study, the scale was reverse-coded in order for the higher value to represent greater satisfaction with life.

of *likes* in a given geographical location.

The data on Facebook *likes* in each ZIP code area were downloaded from the New York Times website, which published a study and an associated interactive map about the distribution of college football fans throughout the USA.[16]

Information on likes for these ZIP codes was then matched to data in specific counties based on the division in the 2010 census. In order to link the ZIP codes to our county-identified observations, we used the 2010 ZIP Code Tabulation Area (ZCTA) Relationship File provided by the US census.[17] Percentages from these ZIP codes were then weighted by their respective populations in order to obtain the relative percentage of *likes* for each applicable county.

In total, the six years of BRFSS surveyed 2,440,925 respondents. After restricting the sample for the period of one week prior to the first and one week after the last game of each season and matching the data to football results, we obtained the dataset of 576,128 observations. However, a substantial issue with this simple matching is that it links all observations in a given area to one team, which may not be actually supported by all football fans living in the area, introducing a measurement error into the model.

In order to mitigate this issue, the sample was further restricted to only take into account areas where a specific team can be considered dominant. Therefore, only areas where the major team claims more than half[18] of the total number of fans are used.[19] Thus, the baseline sample includes 176,262 observations.

Although this reduces the sample size, this step should arguably help to reveal the effect in question. Nevertheless, given that it is impossible to directly identify whether the particular respondent is a football fan or not, our empirical analysis will produce intention-to-treat (ITT) estimates. Hence, the estimated effect will likely

[16] http://www.nytimes.com/interactive/2014/10/03/upshot/ncaa-football-map.html

[17] https://www.census.gov/geo/maps-data/data/zcta_rel_layout.html

[18] Note that the actual choice of cutoff percentage does not substantially alter the results (see Section 1.6.1).

[19] In the hypothetical case of a county where one team had 51% of fans and the second team had 49%, our methodology would not be able to capture the dominant team. However, this is not the case in our data. The smallest difference between the top two teams is 16 percentage points (51% vs. 36%) and only about 5% of the survey responses come from counties with a difference below 30 percentage points. Excluding areas with relatively smaller percentage difference from the estimation does not qualitatively alter the results.

be biased downward.

The specific frequencies of the life satisfaction categories in the baseline sample are reported in Table 1.3.2. Note that the vast majority of responses falls into the

Table 1.3.2: Life Satisfaction Frequencies

Life Satisfaction	No.	Col %	Cum %
Very Dissatisfied	2,082	1.2	1.2
Dissatisfied	8,313	4.7	5.9
Satisfied	86,860	49.3	55.2
Very Satisfied	79,007	44.8	100.0
Total	176,262	100.0	

Source: BRFSS for period from 2005 to 2010.
Area coverage shown in Figure 1.3.3.

top two of the four categories, which complicates the analysis because smoother adjustments along the scale are not possible. However, as larger changes in the valuation of life satisfaction are needed to prospect into its measurement, this could be viewed as a type of attenuation bias in the sense that some information is lost by rounding of the actual feeling.[20]

Areas included in the analysis are depicted in Figure 1.3.3. Examining the composition of teams in the data,[21] the University of Oregon and Louisville are the only two teams that have majority support from outside their state borders. Moreover, states that are generally strong in football such as Texas, California, and Alabama contain counties with differing team fan bases within the state.

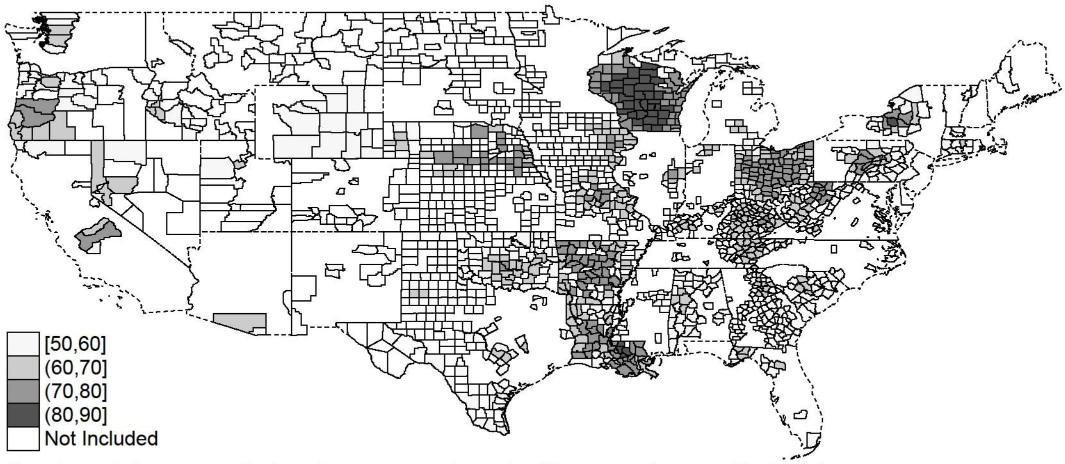
1.4 Methodology

As our dependent variable, life satisfaction, is measured on an ordinal scale, a limited dependent variable model was used. Specifically, an ordered logit model was selected, as its functional form allows for fixed effects.

[20]Statistically, while it increases the chance of a type II error, it decreases the chance of a type I error.

[21] For a complete list of teams and states, see Table 1.B.1 in the Appendix.

Figure 1.3.3: Areas Included in the Analysis



Note: Legend shows categories based on percentage intervals of fans supporting a specific football team.

The functional form of the model follows the equation

$$y^*_{ijt} = \theta_j + \xi_t + X_{ijt}\beta + g\left(S_{jt}, w_{jt}, d_{ijt}\right) + \varepsilon_{ijt} \tag{1.1}$$

$$y_{ijt} = k \qquad \text{if } \kappa_{k-1} < y^*_{ijt} \leq \kappa_k \tag{1.2}$$

where θ_j and ξ_t are regional and time fixed effects and X_{ijt} is a vector of control variables described below. The function $g\left(S_{jt}, w_{jt}, d_{ijt}\right)$ was designed to capture the effects of football results and their (un)expectedness. It takes the form

$$\begin{aligned} g\left(S_{jt}, w_{jt}, d_{ijt}\right) =& \lambda_1 \cdot 1\left[S_{jt} \geq 9\right] \cdot 1\left[w_{jt} = 1\right] \cdot 1\left[0 < d_{ijt} \leq 3\right] + \\ & \lambda_2 \cdot 1\left[S_{jt} \leq -9\right] \cdot 1\left[w_{jt} = 0\right] \cdot 1\left[0 < d_{ijt} \leq 3\right] + \\ & \gamma_1 \cdot 1\left[S_{jt} \geq 9\right] \cdot 1\left[w_{jt} = 1\right] + \\ & \gamma_2 \cdot 1\left[S_{jt} \leq -9\right] \cdot \left[w_{jt} = 0\right] + \\ & \delta_1 \cdot 1\left[w_{jt} = 1\right] + \\ & \delta_2 \cdot 1\left[0 < d_{ijt} \leq 3\right] + \\ & \delta_3 \cdot 1\left[w_{jt} = 1\right] \cdot 1\left[0 < d_{ijt} \leq 3\right], \end{aligned} \tag{1.3}$$

where S_{jt} denotes the pre-game betting spread, w_{jt} is a dummy variable equal to one if the specific team won the previous game and zero if it lost, and d_{ijt} is the number of days between the previous game and date of the survey, indicating whether the game fell into the *post-game window*, defined as within the period of three days after the particular game was played.

The selection of the length of the post-game window lies mainly in the fact that the sample of observations in periods where teams play week-by-week games is broken down to approximately half of the period between the two games. We suspect that the effect would be stronger within a shorter period. However, we decided to choose a relatively longer period in order to ensure a sufficient number of identifying observations (note that as we do not know the exact timing of the survey, we need to exclude days when a game took place). The robustness of this selection is presented in Section 1.6.1.

Our particular research interest lies in parameters λ_1 and λ_2. Specifically, if only unexpected football results during the post-game window have effects on the life satisfaction of the population, λ_1 would be positive, and λ_2 would be negative, while the other coefficients of $g(S_{jt}, w_{jt}, d_{ijt})$ would be zero. If unexpected results have effects regardless of whether the survey takes place in the post-game window, coefficient γ_1 would be positive and coefficient γ_2 negative. If there is an effect of a win in the post-game window in general, but there is no additional effect of an unexpected win, coefficient δ_3 would be positive along with λ_1 and γ_1 being zero.

Coefficient δ_1 measures the general effect of a win, coefficient δ_2 controls for a potential effect of the post-game window, and coefficient δ_3 measures a general effect of a win in the post-game window.

Based on the results of previous studies (see e.g. Dolan et al. 2008) and on the data available, the control variables contained in vector X_{ijt} can be broken down into several categories. First, we include the data on an individual's characteristics - age and age squared, gender, and whether there are children living in the household. Second, we include several sets of dummies reflecting the respondents' marital status, employment status, education, and income. Third, health proxies are included - variables on participation in physical exercise, being limited in activity and variables regarding smoking are used. See Table 1.A.1 in the Appendix for an overview of survey questions associated with these variables.

1.5 Results

1.5.1 Baseline Analysis

Results of the analysis are presented in Tables 1.5.1 and 1.5.2, with the former showing several regressions, including the baseline in column 5, and the latter presenting probability derivatives from the baseline regression. Standard errors in all regressions were adjusted for clustering at the county level,[22] and estimations starting with the fourth column include the set of football variables, the vector of controls, weekly fixed effects, and team-state fixed effects.

Note that, with the exception of Tables 1.5.2 and 1.6.2, all regression-related tables in this study present regression coefficients rather than marginal effects. This is because with the four outcomes of the dependent variable, the ordered logit model implies four different marginal effects, which would make the outputs of our regressions much less tractable.

We can see that the coefficient on an unexpected win in the post-game window is positive and statistically significant throughout all specifications. However, coefficients for an unexpected loss remain insignificant in all regressions. These findings suggest that the effects of *unexpected* win and loss are not symmetrical. In this sense, the results present field evidence of the existence of the reference dependent preferences of Koszegi and Rabin (2006).

The fact that an effect is found for unexpected wins but not losses at first seems surprising in view of previous knowledge. As noted earlier, Card and Dahl (2011) found an increase in family violence following an unexpected loss, but no decrease after an unexpected win. Similarly, Eren and Mocan (2018) found that unexpected losses by the LSU football team lead to an increased length of juvenile sentences given out by judges who received their bachelors degrees from LSU, while unexpected wins do not lead to shorter sentences.

While our findings may at first seem to contradict Card and Dahl's (2011) and

[22]The standard errors from the baseline case of clustering on the county level only change negligibly when the model is estimated with clustering at the weekly level, Huber/White heteroscedasticity consistent estimator, or without adjusting the standard errors.

Table 1.5.1: Baseline Regression: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	(1)	(2)	(3)	(4)	(5)	(6)
λ_1: Unexp. Win[1] $\times$ Post-Game[2]	.203***	.248***	.24***	.239***	.541***	.136*
	(0.07)	(0.08)	(0.08)	(0.08)	(0.16)	(0.08)
λ_2: Unexp. Loss[1] $\times$ Post-Game[2]	.038	.016	1.6e-04	9.5e-03	-.025	.172
	(0.06)	(0.06)	(0.06)	(0.07)	(0.09)	(0.12)
γ_1: Unexpected Win[1]	-.043	-.104**	-.089**	-.076*	-.243*	-.014
	(0.04)	(0.04)	(0.04)	(0.05)	(0.13)	(0.05)
γ_2: Unexpected Loss[1]	.033	.033	.042	.025	.03	3.5e-03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.08)
δ_1: Win	-.017	-9.1e-03	-9.1e-03	-.014	3.1e-03	-.04*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
δ_2: Post-Game Window[2]	.011	-.021	-.016	-3.8e-03	-.023	9.4e-03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
δ_3: Win $\times$ Post-Game Window[2]	-.014	-7.2e-05	-2.1e-04	-.015	-.018	5.6e-03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)
Controls[3]	No	Yes	Yes	Yes	Yes	Yes
Weekly fixed effects	No	No	Yes	Yes	Yes	Yes
State-team fixed effects	No	No	No	Yes	Yes	Yes
Observations	176,262	173,431	173,431	173,431	84,470	88,961
Games Included	All	All	All	All	Home	Road

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
[3] Controls include an individual's personal, economic and health variables. See Appendix 1.A and the supplementary material for details.
Source: Estimation of the ordered logit model.

Eren and Mocan's (2018) studies, they in fact present complementary rather than substitute evidence. Due to the nature of the dataset discussed in Section 1.3.3 (see also Table 1.3.2), our methodology is more sensitive to positive changes in SWB and very likely insensitive to its small negative changes. This is in contrast with Card and Dahl's (2011) and Eren and Mocan's (2018) approach; they observe outcomes associated with negative SWB, arguably making their methodology insensitive to small positive SWB changes.

The combined interpretation of our and previous results is that unexpected football results may likely affect outcomes in both positive and negative domains of SWB, depending on the prevailing emotional aspect connected to the specific outcome. In other words, unexpected wins are likely to affect variables linked to positive emotions such as life satisfaction, and unexpected losses are likely to affect negative phenomena such as domestic violence (Card and Dahl, 2011) or disposition lengths (Eren and Mocan, 2018). In any case, the opposite football outcome, even if unexpected, does not seem to effect the outcome at hand. Note that this implies that even though every unexpected win of one team inevitably carries an unexpected loss of the team's opponent, emotional shocks caused by unexpected football results do not form a zero-sum game.

The regressions based on the sample broken by whether the game was played at home or on the road are presented in columns (5) and (6). The results reveal that the overall effect is driven predominantly by home games. Our interpretation of this fact is that the social context of experiencing the wins with other likewise minded individuals is the driving factor for this result. Because the evidence suggests that home games seem to matter, results in the remainder of the study concentrate on the home-game effects in particular, with the regression in column (5) as the baseline specification. Results of the analysis on the full sample are available upon request.

The marginal effects from the baseline estimation are shown in Table 1.5.2. Following an unexpected win, the probability that a respondent reports being *very satisfied* rises by almost 12 percentage points regardless of which combinations of football covariates one considers. This suggests that, on average, every ninth person would overestimate their actual life satisfaction following an unexpected win. Note that, as discussed in Section 1.3.3, this 12% is an intention-to-treat (ITT) estimate, and is thus likely downward-biased.

Table 1.5.2: Baseline Regression: Marginal Effects of Unexpected Wins

Life Satisfaction	Probability[1]		Post-Game[2]		Outside[3]	
	Sample	Model	ME	(Low, High)	ME	(Low, High)
Very Dissatisfied	0.012	0.012	-0.006	(-0.010, -0.003)	-0.006	(-0.010, -0.002)
Dissatisfied	0.047	0.048	-0.023	(-0.036, -0.010)	-0.022	(-0.035, -0.009)
Satisfied	0.490	0.492	-0.088	(-0.139, -0.036)	-0.089	(-0.141, -0.037)
Very Satisfied	0.451	0.448	0.117	(0.048, 0.185)	0.117	(0.048, 0.186)

Table shows the marginal effects of an unexpected win in the post-game window (λ_1).
All coefficients are statistically significant at 99%.
95% confidence intervals reported in parenthesis.
[1] Probability of the survey answer to the life satisfaction question in the estimation sample and predicted probability of the particular answer from the estimated model.
[2] Marginal effect of an unexpected win in the post-game window compared to a general win in the post-game window.
[3] Marginal effect of an unexpected win in the post-game window compared to a general win outside of the post-game window.
Source: Estimation of the ordered logit model.

In order to estimate the possible size of the effect, we can use a simple back-of-the-envelope calculation. Given that market research estimated 43% of US citizens were college football fans in 2012[23] and assuming the distribution of fan percentages to be homogeneous across the United States, the rescaled effect would be approximately 27 percentage points. However, note that our estimation only includes regions with high fan support for one team. It is not unlikely that such regions will also have a higher overall share of fans in the population, which would in turn bias our back-of-the-envelope estimate upwards. Thus, we can conclude that the true size of the effect lies somewhere between 12 and 27 percentage points.

Even though the effect is statistically significant and may be seen as sizable, it may also be viewed as negligible from the point of view of the overall aggregated measure. Specifically, the data show that the long-term mean is distorted by a fraction of 0.0004 of a standard deviation in the overall data set. This means that the effect does not present an issue for life satisfaction comparisons through regions and/or time.

In terms of a policy application, our results do not bring good news for advocates

[23] http://sportsaffiliates.learfieldsports.com/files/2012/11/College-vs.-Pro.pdf

of stadium subsidies. While economists generally agree that sports events and stadiums do not carry measurable economic benefits to the particular regions (Coates and Humphreys, 2008; Baade et al., 2008), a recent conjecture is that such subsidies could be supported by the fact that sports events bring a certain "feel-good" factor (see Section 1.2.2). Our results indicate that only unexpected wins generate increased life satisfaction. Hence, increases in subjective well-being cannot justify such subsidies.

Coefficients on most of the control variables are strongly statistically significant with a sign that is in line with the previous literature.[24] However, as this study concentrates on the effects of football on life satisfaction, the coefficients of these control variables are not reported here. Full regression results are presented in Appendix 1.D.

1.5.2 Social Context: Sharing the Wins Together

The finding that the effects of football on life satisfaction are driven by games played at the home stadium indicates that the social context of experiencing the win with other like-minded individuals may be the underlying driving factor behind the results. If that is the case, areas with relatively larger football stadiums should report stronger effects.

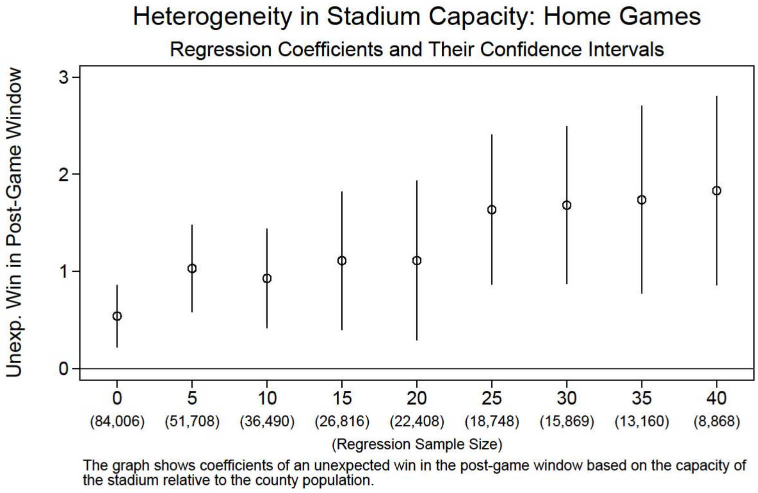
In order to examine this mechanism, we calculated the relative stadium size as the ratio of the stadium capacity and the population of the county where the stadium lies. The baseline regression was then reestimated to include only areas where the relative stadium size is at least as high as some specific percentage.

The coefficients of an unexpected win in the post-game window based on the minimum required stadium size are shown in Figure 1.5.1. The fact that these effects increase with the stadium size relative to the local population suggest that the social context is likely the driving factor.

Note, however, that stadium capacity may proxy for the general importance of the football team to the local community. Therefore, if normalized for the county

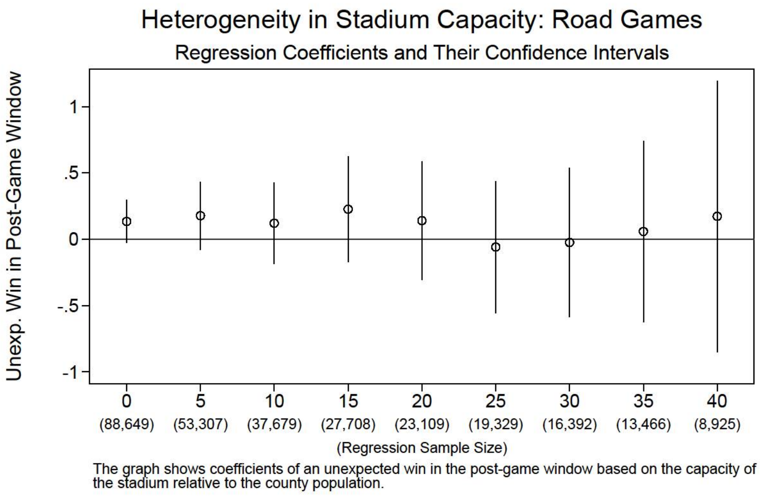
[24] For example, life satisfaction follows a U-shaped pattern throughout individuals' age (Blanchflower and Oswald, 2008), household income generally has a positive effect (Huang and Humphreys, 2012), and children seem to be associated with lower life satisfaction (Deaton and Stone, 2014).

Figure 1.5.1: Stadium Capacity Relative to County Population



population, it serves as an indicator of how important the specific football team is in the local society. If the increased general team support rather than the social context was the main reason for the increasing effects in Figure 1.5.1, the effect would be upward sloping when unexpected wins occur in road games as well. However, as can be seen from Figure 1.5.2, this is clearly not the case. Thus, the evidence is consistent with the social context being the likely reason.

Figure 1.5.2: Stadium Capacity Relative to County Population



Finally, because home games are often associated with substantial consumption of alcohol,[25] there is a possibility that the effect is driven by alcohol consumption

[25]Lindo et al. (2018) use the timing of football games to establish a link between alcohol con-

rather than by the social context. However, due to our comparison of unexpected and general football results, and the fact that most of the alcohol consumption arguably takes place before and during the game, the alcohol consumption levels should be similar in the control and treatment groups. Indeed, examination of the dataset reveals that there is no substantial difference between alcohol-related survey responses based on game outcomes. Hence, alcohol is unlikely to be the reason behind our results.

1.5.3 Demographic Effect Heterogeneity

The previous sections suggest that unexpected wins by home teams have positive effects on the life satisfaction of residents in the locality of the team. However, the possibility of this effect being heterogeneous between demographic groups has not been addressed. In this section, we utilize the advantage of a relatively large sample and attempt to identify demographic groups for which the effect may differ.

Personal Characteristics

This section presents results of regressions on subsamples based on gender and education. The education-based distinction is important due to the fact that the study analyzes results of college teams - while non-graduates may still identify with a college team, the effects should arguably be stronger for alumni.

The results are reported in Table 1.5.3. Note that, in all the tables remaining in the main body of the manuscript, only the coefficients λ_1 and λ_2 are reported. Results including all football-related covariates can be found in Appendix 1.C.

As expected, the point estimate of the effect for college graduates is larger than for non-graduates. This is likely because being a college alumni creates a psychological attachment to the school; hence, the emotions and feelings related to the particular football team may likely be stronger. However, note that the two coefficients are not statistically different from each other.

Interestingly, there seems to be no difference in effects on female and male respondents. We find this result interesting as men are generally viewed as being

sumption and rape.

associated with stronger fan connections than women.

Table 1.5.3: Breakdown by Demographic Groups: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	Gender		College Graduate		
	M	F	Yes	No	All
λ_1: Unexp. Win[1] × Post-Game[2]	.528***	.542**	.677**	.479**	.541***
	(0.18)	(0.24)	(0.27)	(0.21)	(0.16)
λ_2: Unexp. Loss[1] × Post-Game[2]	.193	-.148	.073	-.069	-.025
	(0.13)	(0.10)	(0.15)	(0.11)	(0.09)
Observations	31425	53045	26555	57915	84470

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All columns include full set of controls, weekly fixed effects, and state-team fixed effects.
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
See Table 1.C.1 for results including all football-related covariates.
Source: Estimation of the ordered logit model.

County Attributes

The second distinction of the effect explores possible heterogeneity based on the geopolitical county position. Specifically, we ran separate regressions where the sample was broken down based on whether the county is in a metropolitan statistical area (MSA), and the political preference of the county's citizens in the 2008 presidential elections.[26] Results of the analysis can be found in Table 1.5.4.

The results indicate that the effect in question may be stronger in non-metropolitan areas. This is not surprising as it could be argued that college football is mainly followed in areas with lower population density. Although the two coefficients are statistically not significantly different, these results may indicate why previous studies on the effects of sports did not concentrate on the relationship between football and life satisfaction, as they mostly used data from MSA areas only.

The second distinction shows the analysis broken down into counties that voted

[26] Although the debate about the polarization of the American electorate is recently livelier than ever, research has shown that election decisions are based on a wider set of domains than purely economical (see introduction to Ansolabehere et al. (2006) for more information). Therefore, there is a chance that attitudes towards sports differ between voters of the two parties.

for Democratic and Republican candidates in the 2008 presidential election.[27] Interestingly, these results suggest that the overall effect is driven by counties with majority support for Republicans. Although there are several possible explanations for this effect, all of them are likely linked by the fact that the demographic characteristics of Republican voters substantially overlap with those of football fans. In fact, a study by the National Media Research, Planning and Placement (NRMPP), analyzing data from 2008 and 2009, has shown that college football is the second most Republican-supported sport, in between PGA golf and Nascar racing.[28] According to this study, college football is followed by mostly Republican fans, while Democratic fans more often follow other sports such as NBA or tennis. In light of this result, it is not surprising that our findings are driven by counties with predominantly Republican support.

Table 1.5.4: Breakdown by County Characteristics: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	MSA[3]		Politics[4]		
	Yes	No	Dem	Rep	All
λ_1: Unexp. Win[1]× Post-Game[2]	.367**	1.09***	.277	1.44***	.541***
	(0.17)	(0.26)	(0.17)	(0.28)	(0.16)
λ_2: Unexp. Loss[1]× Post-Game[2]	-.02	-.025	.097	-.138	-.025
	(0.11)	(0.17)	(0.13)	(0.14)	(0.09)
Observations	57645	26825	34986	38401	84470

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All columns include full set of controls, weekly fixed effects, and state-team fixed effects.
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
[3] Coded as "Yes" if the county falls into a Metropolitan Statistical Area.
[4] Counties divided based on results of the 2008 presidential elections. Samples restricted based on having a minimum 5% margin in the final outcome.
See Table 1.C.2 for results including all football-related covariates.
Source: Estimation of the ordered logit model.

[27] To avoid any possible influence of counties that almost tied, we excluded counties where the winning candidate had a margin of less than 5%. Therefore, the sample sizes do not add up to the overall number of observations.

[28] Accessed through wayback machine at `https://web.archive.org/web/20110304071230/http://nmrpp.com/assets/NMRPPsportspolitics.pdf`

1.6 Robustness and Sensitivity Checks

This section presents results of several types of robustness and sensitivity analysis. We begin with an exploration of the cutoff values that were selected for the baseline estimation. The second subsection then proceeds with a discussion of the composition of the control group, functional form specification, and placebo tests.

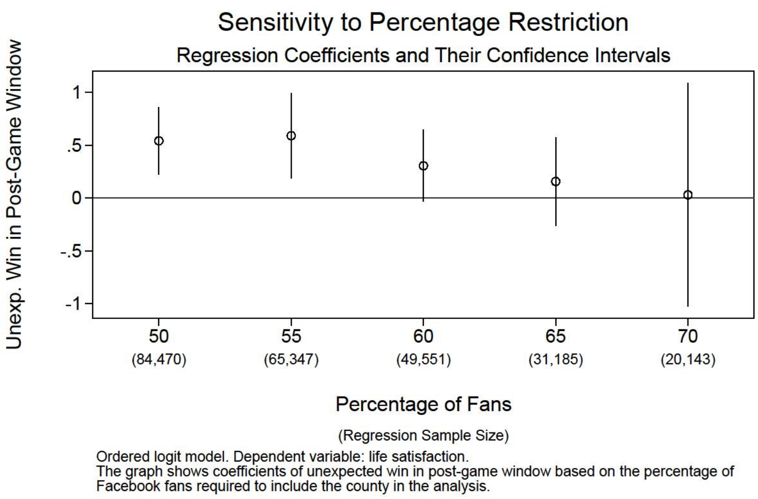
1.6.1 Selection of Cutoff Values

The results of several robustness checks on the coefficient of unexpected wins in the post-game window are presented in following sections. Due to space constraints, results in this section are presented graphically. Full scale tables reporting estimates from these regressions are space-demanding and are available upon request. The controls maintain their approximate significance levels throughout all robust estimations. The coefficient on unexpected losses remains insignificant.

Sample Restriction Based on Like Percentage

The results covering the sensitivity of our baseline regression to the selection of the cutoff percentage rate for sample restriction are presented in Figure 1.6.1.

Figure 1.6.1: Sensitivity to Percentage of Likes

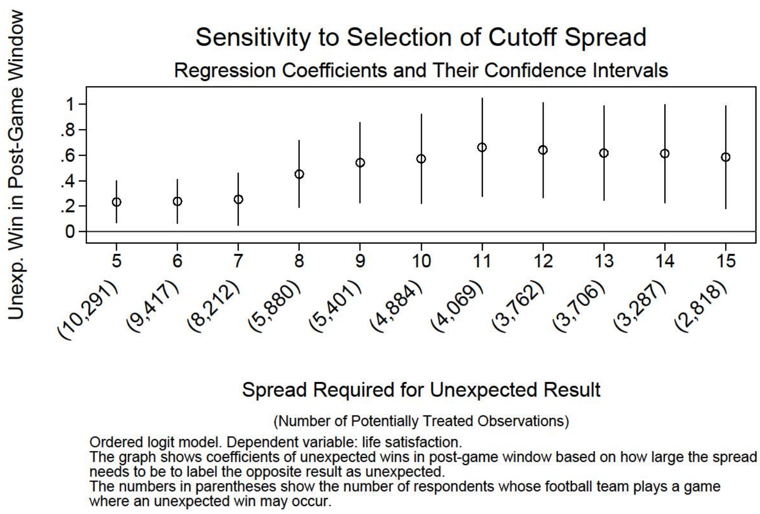


We can see that increasing the cutoff rate generally leads to a higher reported point estimate, suggesting the idea that the effect is stronger in areas where the dominant team has higher support. If, however, it reaches an area above 65%, the number of observations declines as the sample size decreases substantially, in turn harming statistical inference and expanding standard errors of regression coefficients.[29]

Point Difference For Unexpected Results

In order to check for potential sensitivity to how unexpected the outcome is, we adapt several changes of the default cutoff. The results are presented in Figure 1.6.2.

Figure 1.6.2: Sensitivity to Value of Spread



The figure suggests that the stronger the surprise is, the stronger the relationship is. Moreover, the coefficient is statistically significant regardless of which value of the cutoff spread is chosen.

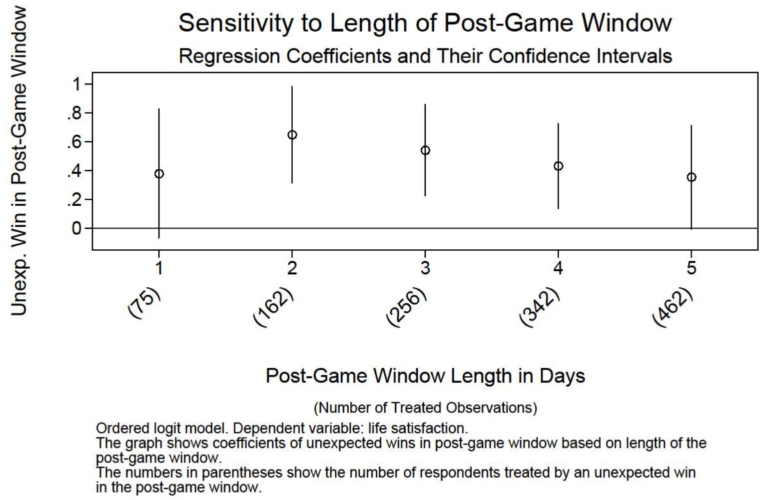
Post-Game Window Length

The results of regressions depending on the length of the post-game window are presented in Figure 1.6.3. The effects for one- and two-day periods are arguably not identified due to a small number of observations in the treatment group (hence the

[29] The largest value of *like* rate is just over 86% of *likes* and only about 10% of observations lies in regions with more than 65% of *likes*.

larger standard error). The results also show that the effect does not disappear even after expanding the post-game window length to five days.

Figure 1.6.3: Sensitivity to Length of Post-game Window



1.6.2 Specifics of Empirical Methodology

This section presents results of three robustness checks which examine potential issues with assumptions behind the empirical methodology employed in the estimation. Namely, the following sections examine the design of the control group, selection of the functional form of the model. The section concludes by description of the placebo test used to validate the results.

Composition of the Control Group

The regression design described in Section 1.4 carries a glitch as the first three days after a game are included in the treatment group, while the fourth and following days enter the control group. Therefore, there is a danger of the benchmark level of life satisfaction being influenced by the treatment variable. In order to examine this concern, estimation of the baseline model was repeated using different sample restrictions based on whether there was a previous unexpected result that could possibly have influenced the control group.

The results are shown in Table 1.6.1. The first column shows results from the baseline estimation and is therefore present for comparison reasons only. The second column excludes games which ended with an unexpected result in the two weeks after the previous unexpected result. The third column excludes all observations that happened after the first unexpected result in a given season. Finally, the fourth column includes only weeks before and after an unexpected result which occurred at least two weeks after the previous unexpected result.

The similarity of all coefficients in these regressions suggests that there is only a very limited methodological concern in terms of control group composition.

Table 1.6.1: Control Group Composition: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	All[A]	Excl Week[B]	Excl All[C]	Incl[D]
λ_1: Unexp. Win[1] × Window[2]	.541***	.599***	.649***	.577***
	(0.16)	(0.15)	(0.16)	(0.22)
λ_2: Unexp. Loss[1] × Window[2]	-.025	-.047	-.092	.046
	(0.09)	(0.09)	(0.10)	(0.18)
Observations	84470	81430	66600	6735

Coefficients from regressions with alternative definition of control groups described below.
All columns include full set of controls, weekly fixed effects, and state-team fixed effects.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[A] Baseline estimation.
[B] Excludes the week after an unexpected result.
[C] Excludes all observations after the first unexpected result in the season.
[D] Includes only the weeks before and after an unexpected result.
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
See Table 1.C.3 for results including all football-related covariates.
Source: Estimation of the ordered logit model.

Functional Form Specification

While previous sections look at the sensitivity of the main analysis in terms of selecting cutoff values that inevitably remain arbitrary, this section leaves these cutoff values at their baseline levels and explores a potential threat of a different kind. Specifically, as the ordered logit model is heavily dependent on its functional form specification, this section runs an alternative version of the analysis.

As shown in Table 1.3.2, 94.1% of answers to the life satisfaction question lie in categories "Satisfied" and "Very Satisfied". This opens a possibility to check for a functional form misspecification as the effect is very likely identified through transition between the top two categories. Therefore, we excluded the observations in which life satisfaction was reported as "Very Dissatisfied" or "Dissatisfied" and then fit a linear probability model on the resulting binary variable equal to 1 for the "Very Satisfied" answer.

Coefficients on variables of interest from this estimation are reported in Table 1.6.2.[30] The results are qualitatively very similar to results of the baseline model; therefore, we can conclude that functional form misspecification does not present a serious threat in our model.

Table 1.6.2: Robustness to Functional Form: LPM Coefficients
Dependent Variable: 1 if Life Satisfaction reported as "Very Satisfied"

	(1)	(2)	(3)	(4)	(5)
λ_1: Unexp. Win[1] $\times$ Post-Game[2]	.043**	.048**	.047**	.045**	.122***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
λ_2: Unexp. Loss[1] $\times$ Post-Game[2]	.016	.012	9.1e-03	.011	4.4e-03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Controls[3]	No	Yes	Yes	Yes	Yes
Weekly fixed effects	No	No	Yes	Yes	Yes
State-team fixed effects	No	No	No	Yes	Yes
Observations	165,867	163,213	163,213	163,213	79,508
Games Included	All	All	All	All	Home

Linear probability model estimation. Dependent variable coded as 1 if life satisfaction answered as "Very satisfied" and 0 as "Satisfied". Answers "Dissatisfied" and "Very Dissatisfied" dropped from the dataset.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Controls include football variables and an individual's personal, economic and health variables.
See Table 1.C.4 for results including all football-related covariates and Table 1.D.2 for all covariates.
Source: Estimation of the ordered linear probability model.

[30] Results including all football variables are presented in Table 1.C.4 in Appendix 1.C. Full results are shown in the supplementary material in Table 1.D.2.

Placebo Games

Even though we performed robustness checks, the individual heterogeneity present due to the repeated cross-sectional nature of the dataset inevitably leads to a danger of biased coefficients. Therefore, following Doerrenberg and Siegloch (2014), we switched all the game results by six months backward and checked whether our regressions still carried their significance.[31] If the results still proved significant, this would suggest that the effect in fact lies in some unobservable factors that were not controlled for in our regression.

Table 1.6.3: Placebo Regression: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	(1)	(2)	(3)	(4)	(5)
λ_1: Unexp. Win[1]$\times$ Post-Game[2]	.024	.056	6.1e-03	-5.6e-03	-.072
	(0.07)	(0.07)	(0.08)	(0.08)	(0.18)
λ_2: Unexp. Loss[1]$\times$ Post-Game[2]	.054	.037	.045	.043	.115
	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)
Controls[3]	No	Yes	Yes	Yes	Yes
Weekly fixed effects	No	No	Yes	Yes	Yes
State-team fixed effects	No	No	No	Yes	Yes
Observations	193,822	191,350	171,236	171,236	82,847
Games Included	All	All	All	All	Home

Dates of all games switched by six months backward to obtain placebo effects.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Controls include football variables and an individual's personal, economic and health variables. See Table 1.C.4 for results including all football-related covariates and Table 1.D.2 for all covariates. See Appendix 1.A and the supplementary material for details.
Source: Estimation of the ordered logit model.

The results of coefficients of interest from these placebo regressions are reported in Table 1.6.3. The disappearance of the effect supports the validity of our results.

[31] Specifically, all games were switched by 26 weeks in order to keep the day of the week identical for all games in question.

1.7 Conclusion

To our knowledge, this study is the first to find statistically significant effects of sports results on life satisfaction in a large scale dataset. Specifically, it presents evidence of an increase in life satisfaction scores following an unexpected win by the local college football team in the three days after the game.

The analysis reveals that the effects are driven entirely by home games, indicating that the social context of a win is the key driving factor behind the effects. Overall, the results present evidence that the impacts of emotional shocks caused by unexpected football wins are larger when the experience is shared with others. This finding is supported by the fact that the size of the effects increases with the size of the team's stadium relative to the local population.

No effects are found for unexpected losses or for results which cannot be viewed as surprising based on the pre-game betting market. As there is no ex ante guarantee that a team will generate any unexpected wins, increases in subjective well-being cannot justify subsidies for sports stadiums.

While the absence of the effect of unexpected losses may at first seem contradictory to the results of previous studies, which found evidence of unexpected losses increasing rates of domestic violence (Card and Dahl, 2011) and judicial sentence lengths (Eren and Mocan, 2018), the difference in the effects is caused by our data likely being more sensitive to small positive changes in SWB and vice-versa.

The identified effect from our baseline regression suggests that the probability of respondents reporting the highest category of life satisfaction rises by approximately 12 percentage points following an unexpected win when surveyed within three days after the game. Moreover, this effect is likely biased towards zero due to the presence of measurement error. A back-of-the-envelope calculation suggests that this effect would be in the range of 12 to 27 percentage points in the absence of this measurement error.

Nevertheless, although the effect is sizable ex-post after an unexpected win, it is important to note that its overall magnitude is negligible. Even though the effect is not zero-sum due to the presence of reactions to unexpected wins but not losses, the overall dataset mean is distorted upwards by a fraction of 0.0004 of the data's standard deviation. Thus, while the effect is statistically significant and precisely estimated, it is too small to present issues for the measure of life satisfaction in the sense of comparing its reported values through time and/or region.

The results are robust to functional form specification, control group definition, restriction on the strength of team support, level of surprise needed in order to designate a result unexpected, and the number of days we consider an individual to be potentially affected by the football game result. After switching dates of games out of the football season in order to test for a placebo effect, the relationship disappears, supporting the existence of the effect.

Note that there is one explanation for the effect that our study was not able to examine. Specifically, the dataset used in the analysis only asks the respondent a question about life satisfaction, without previously examining her momentary happiness. This causes a danger of misreporting life satisfaction, in the sense of respondents being unaware that their current mood may alter their answer. Future research is needed to disentangle these two possibilities.

Appendix

1.A Appendix 1.A: Composition of Control Variables

This section of the Appendix provides the description of control variables, including their respective survey questions and their descriptive statistics.

Table 1.A.1: Description of Dummies from BRFFS Variables

Variable	Survey Question	Coded as 1 if
Children in household	How many children less than 18 years of age live in your household?	there is at least one child
Marital status dummies	Are you: (marital status)	Answer reflects the dummy
Employment status dummies	Are you currently: (employment status)	"
Education dummies	What is the highest grade or year of school you completed?	"
Income dummies	Is your annual household income from all sources:	" (plus missing)

Table 1.A.1: Description of Dummies from BRFFS Variables (continued)

Variable	Survey Question	Coded as 1 if
Physically exercising	During the past month, other than your regular job, did you participate in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise?	"yes"
Limited in activity	Are you limited in any way in any activities because of physical, mental, or emotional problems?	"yes"
Smoking dummies	Do you now smoke cigarettes every day, some days, or not at all?	Answer reflects the dummy (plus missing)

Source: BRFSS and own calculation

Table 1.A.2: Descriptive Statistics of BRFSS Data

	Mean	S.D.
Life Satisfaction	3.378	0.632
Personal demographics		
Age in years	54.894	16.691
Age in years (squared)	3291.945	1843.016
Male	0.374	0.484
Children in household	0.293	0.455
Marital status dummies (baseline: Never married)		
Married	0.569	0.495
Divorced	0.142	0.349
Widowed	0.139	0.346
Separated	0.021	0.142
A member of an unmarried couple	0.021	0.143
Employment status dummies (baseline: Employed for wages)		
Self-employed	0.080	0.271
Out of work for more than 1 year	0.020	0.141
Out of work for less than 1 year	0.024	0.154
Homemaker	0.082	0.274
Student	0.017	0.131
Retired	0.275	0.446
Unable to work	0.072	0.259
Education dummies (baseline: High school graduate)		
Never attended school or only kindergarten	0.001	0.037
Grades 1 - 8 (Elementary)	0.030	0.172
Grades 9 - 11 (Some high school)	0.066	0.249
College 1 to 3 years (Some college or technical school)	0.270	0.444
College 4 years or more (College graduate)	0.316	0.465
Income dummies (baseline: $35,000 to under $50,000)		
Annual household income under $10,000	0.046	0.210
Annual household income $10,000 to under $15,000	0.055	0.227
Annual household income $15,000 to under $20,000	0.070	0.255
Annual household income $20,000 to under $25,000	0.091	0.287
Annual household income $25,000 to under $35,000	0.116	0.320
Annual household income $50,000 to under $75,000	0.144	0.351
Annual household income over $75,000	0.209	0.406
Income info missing	0.127	0.333
Health proxies		
Physically exercising	0.720	0.449
Limited in activity	0.267	0.442
Smoking every day	0.138	0.345
Smoking some days	0.044	0.204
Quit smoking	0.290	0.454

Source: BRFSS

1.B Appendix 1.B: Teams Used in the Analysis

Table 1.B.1: Teams and States in the Analysis

	Counties	Observations
Alabama (Alabama)	30	4330
Arizona (Arizona)	2	1614
Arkansas (Arkansas)	55	6592
Auburn (Alabama)	2	193
Boise State (Idaho)	9	3091
Connecticut (Connecticut)	6	7619
Florida State (Florida)	5	1704
Florida (Florida)	4	1762
Fresno State (California)	3	527
Georgia (Georgia)	45	1915
Illinois (Illinois)	7	290
Iowa State (Iowa)	2	276
Iowa (Iowa)	34	4103
Kansas (Kansas)	1	566
Kentucky (Kentucky)	62	6362
Louisville (Indiana)	3	267
Louisville (Kentucky)	3	1120
LSU (Louisiana)	55	10083
Miami (Florida)	1	776
Michigan State (Michigan)	4	607
Michigan (Michigan)	3	721
Mississippi State (Mississippi)	4	544
Missouri (Missouri)	43	3917
North Carolina (North Carolina)	4	1110
Nebraska (Nebraska)	39	12576
Nevada (Nevada)	6	3606
New Mexico (New Mexico)	2	1953
Notre Dame (Indiana)	4	867
Ohio State (Ohio)	86	14085

Source: Author's calculation

(Continued on the next page)

Table 1.B.1: Teams and States in the Analysis (continued)

	Counties	Observations
Oklahoma State (Oklahoma)	1	205
Oklahoma (Oklahoma)	47	12776
Oregon State (Oregon)	1	237
Oregon (California)	2	79
Oregon (Oregon)	20	8994
Penn State (Pennsylvania)	33	8764
Purdue (Indiana)	2	216
South Carolina (South Carolina)	20	4991
Syracuse (New York)	17	1391
Tennessee (Tennessee)	31	2762
Texas A&M (Texas)	2	67
Texas Tech (Texas)	4	912
Texas (Texas)	21	3899
Utah (Utah)	2	4024
Virginia Tech (Virginia)	6	176
Washington (Washington)	8	12828
West Virginia (West Virginia)	35	4617
Wisconsin (Wisconsin)	67	8775
Wyoming (Wyoming)	13	4542

Source: Author's calculation

1.C Appendix 1.C: Tables with All Football Covariates

In order to save space in the main body of the text, the regression tables except for Table 1.5.1 include only the main coefficients of interest λ_1 and λ_2. This appendix contains the same regression tables including all of the football-related coefficients.

Table 1.C.1: Breakdown by Demographic Groups: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	Gender		College Graduate		
	M	F	Yes	No	All
λ_1: Unexp. Win[1]× Post-Game[2]	.528***	.542**	.677**	.479**	.541***
	(0.18)	(0.24)	(0.27)	(0.21)	(0.16)
λ_2: Unexp. Loss[1]× Post-Game[2]	.193	-.148	.073	-.069	-.025
	(0.13)	(0.10)	(0.15)	(0.11)	(0.09)
γ_1: Unexpected Win[1]	-.254	-.222*	-.317	-.203	-.243*
	(0.22)	(0.13)	(0.21)	(0.13)	(0.13)
γ_2: Unexpected Loss[1]	-.031	.07	.087	7.1e-03	.03
	(0.09)	(0.08)	(0.11)	(0.07)	(0.06)
δ_1: Win	-.029	.029	.024	-3.0e-03	3.1e-03
	(0.04)	(0.03)	(0.05)	(0.03)	(0.03)
δ_2: Post-Game Window[2]	-.052	-3.3e-03	-9.3e-03	-.027	-.023
	(0.05)	(0.04)	(0.06)	(0.04)	(0.03)
δ_3: Win × Post-Game Window[2]	-.024	-.016	-.091	7.1e-03	-.018
	(0.06)	(0.04)	(0.06)	(0.05)	(0.04)
Observations	31425	53045	26555	57915	84470

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All columns include full set of controls, weekly fixed effects, and state-team fixed effects.
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
Source: Estimation of the ordered logit model.

Table 1.C.2: Breakdown by County Characteristics: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	MSA[3]		Politics[4]		
	Yes	No	Dem	Rep	All
λ_1: Unexp. Win[1] $\times$ Post-Game[2]	.367**	1.09***	.277	1.44***	.541***
	(0.17)	(0.26)	(0.17)	(0.28)	(0.16)
λ_2: Unexp. Loss[1] $\times$ Post-Game[2]	-.02	-.025	.097	-.138	-.025
	(0.11)	(0.17)	(0.13)	(0.14)	(0.09)
γ_1: Unexpected Win[1]	-.245*	-.223	-.262	-.403*	-.243*
	(0.15)	(0.24)	(0.16)	(0.23)	(0.13)
γ_2: Unexpected Loss[1]	.027	.044	.01	.039	.03
	(0.07)	(0.11)	(0.09)	(0.09)	(0.06)
δ_1: Win	-8.9e-03	.019	-7.5e-03	.074*	3.1e-03
	(0.03)	(0.05)	(0.04)	(0.04)	(0.03)
δ_2: Post-Game Window[2]	-8.8e-03	-.065	-.054	.084	-.023
	(0.04)	(0.06)	(0.05)	(0.06)	(0.03)
δ_3: Win $\times$ Post-Game Window[2]	-.043	.038	.02	-.145**	-.018
	(0.05)	(0.07)	(0.06)	(0.06)	(0.04)
Observations	57645	26825	34986	38401	84470

Samples in columns 3 and 4 are restricted based on having a minimum 5% margin in the 2008 presidential elections.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All columns include full set of controls, weekly fixed effects, and state-team fixed effects.
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
[3] Coded as "Yes" if the county falls into a Metropolitan Statistical Area.
[4] Counties divided based on results of the 2008 presidential elections. Samples restricted based on having a minimum 5% margin in the final outcome.
Source: Estimation of the ordered logit model.

Table 1.C.3: Control Group Composition: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	All[A]	Excl Week[B]	Excl All[C]	Incl[D]
λ_1: Unexp. Win[1] × Window[2]	.541***	.599***	.649***	.577***
	(0.16)	(0.15)	(0.16)	(0.22)
λ_2: Unexp. Loss[1] × Window[2]	-.025	-.047	-.092	.046
	(0.09)	(0.09)	(0.10)	(0.18)
γ_1: Unexpected Win[1]	-.243*	-.282**	-.386**	-.137
	(0.13)	(0.13)	(0.15)	(0.21)
γ_2: Unexpected Loss[1]	.03	.028	.053	-.115
	(0.06)	(0.06)	(0.07)	(0.17)
δ_1: Win	3.1e-03	2.9e-03	.015	-.207
	(0.03)	(0.03)	(0.03)	(0.16)
δ_2: Post-Game Window[2]	-.023	-.016	.021	-.142
	(0.03)	(0.03)	(0.04)	(0.15)
δ_3: Win × Post-Game Window[2]	-.018	-.026	-.058	.038
	(0.04)	(0.04)	(0.04)	(0.19)
Observations	84470	81430	66600	6735

Coefficients from regressions with alternative definition of control groups described below.
All columns include full set of controls, weekly fixed effects, and state-team fixed effects.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[A] Baseline estimation.
[B] Excludes the week after an unexpected result.
[C] Excludes all observations after the first unexpected result in the season.
[D] Includes only the weeks before and after an unexpected result.
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] Post-game window is a period of three days after the last game was played.
Source: Estimation of the ordered linear probability model.

Table 1.C.4: Robustness to Functional Form: LPM Coefficients
Dependent Variable: 1 if Life Satisfaction reported as "Very Satisfied"

	(1)	(2)	(3)	(4)	(5)
λ_1: Unexp. Win[1]× Post-Game[2]	.043**	.048**	.047**	.045**	.122***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
λ_2: Unexp. Loss[1]× Post-Game[2]	.016	.012	9.1e-03	.011	4.4e-03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
γ_1: Unexpected Win[1]	-.011	-.021**	-.02*	-.015	-.051*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
γ_2: Unexpected Loss[1]	5.9e-03	6.1e-03	6.0e-03	2.4e-03	5.7e-04
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
δ_1: Win	-5.2e-03	-2.5e-03	-2.2e-03	-3.9e-03	-4.0e-04
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
δ_2: Post-Game Window[2]	1.4e-03	-3.6e-03	-2.8e-03	-3.0e-04	-5.2e-03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
δ_3: Win × Post-Game Window[2]	-1.5e-03	-4.4e-05	-4.4e-04	-3.8e-03	-4.6e-03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Controls[3]	No	Yes	Yes	Yes	Yes
Weekly fixed effects	No	No	Yes	Yes	Yes
State-team fixed effects	No	No	No	Yes	Yes
Observations	165,867	163,213	163,213	163,213	79,508
Games Included	All	All	All	All	Home

Linear probability model estimation. Dependent variable coded as 1 if life satisfaction answered as "Very satisfied" and 0 as "Satisfied". Answers "Dissatisfied" and "Very Dissatisfied" dropped from the dataset.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Controls include an individual's personal, economic and health variables. See Appendix 1.A and the supplementary material for details.
Source: Estimation of the ordered linear probability model.

Table 1.C.5: Placebo Regression: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	(1)	(2)	(3)	(4)	(5)
λ_1: Unexp. Win[1]$\times$ Post-Game[2]	.024	.056	6.1e-03	-5.6e-03	-.072
	(0.07)	(0.07)	(0.08)	(0.08)	(0.18)
λ_2: Unexp. Loss[1]$\times$ Post-Game[2]	.054	.037	.045	.043	.115
	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)
γ_1: Unexpected Win[1]	.021	-.014	.031	.055	.066
	(0.03)	(0.04)	(0.04)	(0.04)	(0.11)
γ_2: Unexpected Loss[1]	-4.0e-03	-7.6e-03	-1.6e-03	-5.6e-03	-.023
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
δ_1: Win	-.022*	-.018	-.02	-.016	-.025
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
δ_2: Post-Game Window[2]	-.015	-.045***	-.049***	-.03*	-.07**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
δ_3: Win $\times$ Post-Game Window[2]	.047**	.062***	.066***	.045**	.069**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Controls[3]	No	Yes	Yes	Yes	Yes
Weekly fixed effects	No	No	Yes	Yes	Yes
State-team fixed effects	No	No	No	Yes	Yes
Observations	193,822	191,350	171,236	171,236	82,847
Games Included	All	All	All	All	Home

Dates of all games switched by six months backward to obtain placebo effects.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Controls include an individual's personal, economic and health variables. See Appendix 1.A and the supplementary material for details.
Source: Estimation of the ordered logit model.

1.D Appendix 1.D: Full Regression Results

The supplementary material provided on the following pages shows full regression results associated with regressions from Sections 1.5.1, 1.6.2, and 1.6.2. With some exceptions, these results are comparable to results found by previous research.

Table 1.D.1: Ordered Logit Coefficients
Dependent variable: Life Satisfaction

	(1)	(2)	(3)	(4)
Football results				
Unexp. Win[1] × Window[2]	.248***	.24***	.239***	.541***
	(0.08)	(0.08)	(0.08)	(0.16)
Unexp. Loss[1] × Window[2]	.016	1.6e-04	9.5e-03	-.025
	(0.06)	(0.06)	(0.07)	(0.09)
Unexpected Win[1]	-.104**	-.089**	-.076*	-.243*
	(0.04)	(0.04)	(0.05)	(0.13)
Unexpected Loss[1]	.033	.042	.025	.03
	(0.04)	(0.04)	(0.04)	(0.06)
Win	-9.1e-03	-9.1e-03	-.014	3.1e-03
	(0.02)	(0.02)	(0.02)	(0.03)
Post-Game Window[2]	-.021	-.016	-3.8e-03	-.023
	(0.02)	(0.02)	(0.02)	(0.03)
Win × Post-Game Window[2]	-7.2e-05	-2.1e-04	-.015	-.018
	(0.02)	(0.02)	(0.02)	(0.04)
Personal demographics				
Age in years	-.018***	-.018***	-.019***	-.02***
	(0.00)	(0.00)	(0.00)	(0.00)
Age in years (squared)	2.2e-04***	2.2e-04***	2.4e-04***	2.4e-04***
	(0.00)	(0.00)	(0.00)	(0.00)
Male	-.132***	-.132***	-.129***	-.126***
	(0.01)	(0.01)	(0.01)	(0.02)
Children in household[3]	-.111***	-.112***	-.11***	-.113***
	(0.01)	(0.01)	(0.01)	(0.02)
Marital status dummies (baseline: Never married)				
Married	.622***	.624***	.614***	.657***
	(0.02)	(0.02)	(0.02)	(0.03)
Divorced	.02	.021	.012	.073**
	(0.02)	(0.02)	(0.02)	(0.03)
Widowed	.11***	.111***	.097***	.146***
	(0.02)	(0.02)	(0.02)	(0.03)
Separated	-.232***	-.234***	-.256***	-.212***
	(0.04)	(0.04)	(0.04)	(0.05)
A member of an unmarried couple	.16***	.161***	.175***	.203***
	(0.04)	(0.04)	(0.04)	(0.05)
Employment status dummies (baseline: Employed for wages)				
Self-employed	.124***	.126***	.123***	.138***
	(0.02)	(0.02)	(0.02)	(0.03)
Out of work for more than 1 year	-.57***	-.572***	-.572***	-.576***
	(0.05)	(0.05)	(0.05)	(0.06)
Out of work for less than 1 year	-.571***	-.57***	-.562***	-.617***
	(0.03)	(0.03)	(0.03)	(0.05)
Homemaker	.136***	.14***	.132***	.141***
	(0.02)	(0.02)	(0.02)	(0.03)
Student	.173***	.172***	.164***	.154***
	(0.04)	(0.04)	(0.04)	(0.06)
Retired	.221***	.224***	.212***	.187***
	(0.02)	(0.02)	(0.02)	(0.03)
Unable to work	-.488***	-.487***	-.506***	-.51***
	(0.03)	(0.03)	(0.03)	(0.04)

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Equal to 1 if there are children living in the household with the respondent.

(Continued on the next page)

Table 1.D.1: Ordered Logit Coefficients (continued)
Dependent Variable: Life Satisfaction

Education dummies (baseline: High school graduate)				
Never attended school or only kindergarten	-.095	-.097	-.108	-.306*
	(0.15)	(0.15)	(0.15)	(0.18)
Grades 1 - 8 (Elementary)	-.088***	-.088***	-.116***	-.133***
	(0.03)	(0.03)	(0.03)	(0.04)
Grades 9 - 11 (Some high school)	-.062***	-.063***	-.082***	-.079***
	(0.02)	(0.02)	(0.02)	(0.03)
College 1 to 3 years (Some college or technical school)	2.5e-03	1.7e-03	6.8e-04	2.4e-03
	(0.01)	(0.01)	(0.01)	(0.02)
College 4 years or more (College graduate)	.155***	.154***	.156***	.159***
	(0.01)	(0.01)	(0.01)	(0.02)
Income dummies (baseline: $35,000 to under $50,000)				
Annual household income under $10,000	-.485***	-.485***	-.506***	-.466***
	(0.03)	(0.03)	(0.03)	(0.05)
Annual household income $10,000 to under $15,000	-.411***	-.41***	-.427***	-.377***
	(0.03)	(0.03)	(0.03)	(0.04)
Annual household income $15,000 to under $20,000	-.299***	-.301***	-.318***	-.305***
	(0.02)	(0.02)	(0.02)	(0.03)
Annual household income $20,000 to under $25,000	-.254***	-.254***	-.261***	-.25***
	(0.02)	(0.02)	(0.02)	(0.03)
Annual household income $25,000 to under $35,000	-.171***	-.171***	-.174***	-.14***
	(0.02)	(0.02)	(0.02)	(0.03)
Annual household income $50,000 to under $75,000	.16***	.159***	.163***	.181***
	(0.02)	(0.02)	(0.02)	(0.02)
Annual household income over $75,000	.463***	.461***	.47***	.464***
	(0.02)	(0.02)	(0.02)	(0.03)
Income info missing	-.01	-.012	-.027	-9.0e-03
	(0.02)	(0.02)	(0.02)	(0.03)
Health proxies (smoking baseline: Never smoked)				
Physically exercising[4]	.361***	.364***	.377***	.362***
	(0.01)	(0.01)	(0.01)	(0.02)
Limited in activity[5]	-.721***	-.721***	-.72***	-.72***
	(0.01)	(0.01)	(0.01)	(0.02)
Smoking every day	-.373***	-.372***	-.373***	-.387***
	(0.02)	(0.02)	(0.02)	(0.02)
Smoking some days	-.275***	-.275***	-.277***	-.279***
	(0.03)	(0.02)	(0.03)	(0.03)
Quit smoking	-.065***	-.064***	-.059***	-.066***
	(0.01)	(0.01)	(0.01)	(0.02)
cut1	-4.73***	-4.78***	-4.89***	-4.35***
	(0.06)	(0.10)	(0.11)	(0.44)
cut2	-2.99***	-3.04***	-3.15***	-2.61***
	(0.06)	(0.10)	(0.11)	(0.44)
cut3	.361***	.313***	.213**	.742*
	(0.06)	(0.10)	(0.11)	(0.44)
Weekly fixed effects	No	Yes	Yes	Yes
Team-State fixed effects	No	No	Yes	Yes
Observations	173,431	173,431	173,431	84,470
Games Included	All	All	All	Home

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[4] Equal to 1 if a person participated in physical exercise outside work in past 30 days.
[5] Equal to 1 if activities were limited due to physical, mental, or emotional problems.

Table 1.D.2: Linear Probability Model Coefficients
Dependent Variable: 1 if Life Satisfaction Reported to be "Very Satisfied"

	(1)	(2)	(3)	(4)
Football results				
Unexp. Win[1] × Window[2]	.048**	.047**	.045**	.122***
	(0.02)	(0.02)	(0.02)	(0.04)
Unexp. Loss[1] × Window[2]	.012	9.1e-03	.011	4.4e-03
	(0.01)	(0.01)	(0.01)	(0.02)
Unexpected Win[1]	-.021**	-.02*	-.015	-.051*
	(0.01)	(0.01)	(0.01)	(0.03)
Unexpected Loss[1]	6.1e-03	6.0e-03	2.4e-03	5.7e-04
	(0.01)	(0.01)	(0.01)	(0.01)
Win	-2.5e-03	-2.2e-03	-3.9e-03	-4.0e-04
	(0.00)	(0.00)	(0.00)	(0.01)
Post-Game Window[2]	-3.6e-03	-2.8e-03	-3.0e-04	-5.2e-03
	(0.00)	(0.00)	(0.00)	(0.01)
Win × Post-Game Window[2]	-4.4e-05	-4.4e-04	-3.8e-03	-4.6e-03
	(0.01)	(0.01)	(0.01)	(0.01)
Personal demographics				
Age in years	-2.6e-03***	-2.6e-03***	-2.9e-03***	-2.9e-03***
	(0.00)	(0.00)	(0.00)	(0.00)
Age in years (squared)	3.2e-05***	3.2e-05***	3.6e-05***	3.5e-05***
	(0.00)	(0.00)	(0.00)	(0.00)
Male	-.027***	-.027***	-.026***	-.028***
	(0.00)	(0.00)	(0.00)	(0.00)
Children in household[3]	-.03***	-.03***	-.029***	-.028***
	(0.00)	(0.00)	(0.00)	(0.00)
Marital status dummies (baseline: Never married)				
Married	.136***	.136***	.135***	.14***
	(0.00)	(0.00)	(0.00)	(0.01)
Divorced	.014***	.014***	.012**	.017**
	(0.01)	(0.01)	(0.00)	(0.01)
Widowed	.02***	.02***	.017***	.028***
	(0.01)	(0.01)	(0.01)	(0.01)
Separated	-.02**	-.021**	-.025***	-.019
	(0.01)	(0.01)	(0.01)	(0.01)
A member of an unmarried couple	.029***	.029***	.032***	.036***
	(0.01)	(0.01)	(0.01)	(0.01)
Employment status dummies (baseline: Employed for wages)				
Self-employed	.032***	.032***	.031***	.035***
	(0.00)	(0.00)	(0.00)	(0.01)
Out of work for more than 1 year	-.052***	-.052***	-.052***	-.046***
	(0.01)	(0.01)	(0.01)	(0.01)
Out of work for less than 1 year	-.074***	-.073***	-.072***	-.081***
	(0.01)	(0.01)	(0.01)	(0.01)
Homemaker	.038***	.039***	.036***	.035***
	(0.00)	(0.00)	(0.00)	(0.01)
Student	.044***	.044***	.042***	.032**
	(0.01)	(0.01)	(0.01)	(0.01)
Retired	.06***	.06***	.057***	.048***
	(0.00)	(0.00)	(0.00)	(0.01)
Unable to work	-3.9e-03	-3.7e-03	-8.1e-03	-.016*
	(0.01)	(0.01)	(0.01)	(0.01)

Note: Dependent variable coded as 1 if life satisfaction answered as "Very satisfied" and 0 as "Satisfied".
Answers "Dissatisfied" and "Very Dissatisfied" dropped from the dataset.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Equal to 1 if there are children living in the household with the respondent.

(Continued on the next page)

Table 1.D.2: Linear Probability Model Coefficients (continued)
Dependent Variable: 1 if Life Satisfaction Reported to be "Very Satisfied"

Education dummies (baseline: High school graduate)				
Never attended school or only kindergarten	-.038	-.04	-.042	-.075*
	(0.03)	(0.03)	(0.03)	(0.04)
Grades 1 - 8 (Elementary)	-.036***	-.036***	-.041***	-.041***
	(0.01)	(0.01)	(0.01)	(0.01)
Grades 9 - 11 (Some high school)	-.02***	-.02***	-.024***	-.025***
	(0.01)	(0.01)	(0.01)	(0.01)
College 1 to 3 years (Some college or technical school)	.012***	.012***	.012***	.011**
	(0.00)	(0.00)	(0.00)	(0.00)
College 4 years or more (College graduate)	.052***	.052***	.052***	.051***
	(0.00)	(0.00)	(0.00)	(0.00)
Income dummies (baseline: $35,000 to under $50,000)				
Annual household income under $10,000	-.039***	-.039***	-.043***	-.038***
	(0.01)	(0.01)	(0.01)	(0.01)
Annual household income $10,000 to under $15,000	-.063***	-.063***	-.066***	-.053***
	(0.01)	(0.01)	(0.01)	(0.01)
Annual household income $15,000 to under $20,000	-.052***	-.053***	-.056***	-.048***
	(0.01)	(0.01)	(0.01)	(0.01)
Annual household income $20,000 to under $25,000	-.05***	-.05***	-.052***	-.048***
	(0.01)	(0.01)	(0.01)	(0.01)
Annual household income $25,000 to under $35,000	-.035***	-.035***	-.036***	-.028***
	(0.00)	(0.00)	(0.00)	(0.01)
Annual household income $50,000 to under $75,000	.036***	.036***	.037***	.043***
	(0.00)	(0.00)	(0.00)	(0.01)
Annual household income over $75,000	.109***	.109***	.11***	.109***
	(0.01)	(0.01)	(0.01)	(0.01)
Income info missing	6.5e-03	6.2e-03	3.0e-03	6.1e-03
	(0.00)	(0.00)	(0.00)	(0.01)
Health proxies (smoking baseline: Never smoked)				
Physically exercising[4]	.065***	.066***	.068***	.066***
	(0.00)	(0.00)	(0.00)	(0.00)
Limited in activity[5]	-.125***	-.125***	-.124***	-.124***
	(0.00)	(0.00)	(0.00)	(0.00)
Smoking every day	-.06***	-.06***	-.06***	-.063***
	(0.00)	(0.00)	(0.00)	(0.00)
Smoking some days	-.054***	-.053***	-.054***	-.063***
	(0.01)	(0.01)	(0.01)	(0.01)
Quit smoking	-.014***	-.014***	-.013***	-.015***
	(0.00)	(0.00)	(0.00)	(0.00)
_cons	.397***	.419***	.43***	.435***
	(0.01)	(0.02)	(0.03)	(0.09)
Weekly fixed effects	No	Yes	Yes	Yes
Team-State fixed effects	No	No	Yes	Yes
Observations	163,213	163,213	163,213	79,508
Games Included	All	All	All	Home

Note: Dependent variable coded 1 if life satisfaction reported as "Very satisfied" and 0 as "Satisfied".
Answers "Dissatisfied" and "Very Dissatisfied" dropped from the dataset.
Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[4] Equal to 1 if a person participated in physical exercise outside work in past 30 days.
[5] Equal to 1 if activities were limited due to physical, mental, or emotional problems.

Table 1.D.3: Placebo Regression: Ordered Logit Coefficients
Dependent Variable: Life Satisfaction

	(1)	(2)	(3)	(4)
Football results				
Unexp. Win[1] × Window[2]	.056	6.1e-03	-5.6e-03	-.072
	(0.07)	(0.08)	(0.08)	(0.18)
Unexp. Loss[1] × Window[2]	.037	.045	.043	.115
	(0.05)	(0.06)	(0.06)	(0.07)
Unexpected Win[1]	-.014	.031	.055	.066
	(0.04)	(0.04)	(0.04)	(0.11)
Unexpected Loss[1]	-7.6e-03	-1.6e-03	-5.6e-03	-.023
	(0.04)	(0.04)	(0.04)	(0.05)
Win	-.018	-.02	-.016	-.025
	(0.02)	(0.02)	(0.01)	(0.02)
Post-Game Window[2]	-.045***	-.049***	-.03*	-.07**
	(0.02)	(0.02)	(0.02)	(0.03)
Win × Post-Game Window[2]	.062***	.066***	.045**	.069**
	(0.02)	(0.02)	(0.02)	(0.03)
Personal demographics				
Age in years	-.02***	-.02***	-.022***	-.024***
	(0.00)	(0.00)	(0.00)	(0.00)
Age in years (squared)	2.5e-04***	2.5e-04***	2.7e-04***	2.9e-04***
	(0.00)	(0.00)	(0.00)	(0.00)
Male	-.15***	-.154***	-.15***	-.155***
	(0.01)	(0.01)	(0.01)	(0.02)
Children in household[3]	-.074***	-.071***	-.07***	-.075***
	(0.01)	(0.01)	(0.01)	(0.02)
Marital status dummies (baseline: Never married)				
Married	.593***	.595***	.589***	.604***
	(0.02)	(0.02)	(0.02)	(0.03)
Divorced	.027	.023	.017	.01
	(0.02)	(0.02)	(0.02)	(0.03)
Widowed	.076***	.065***	.054**	.067**
	(0.02)	(0.02)	(0.02)	(0.03)
Separated	-.239***	-.24***	-.262***	-.193***
	(0.04)	(0.04)	(0.04)	(0.06)
A member of an unmarried couple	.246***	.246***	.26***	.271***
	(0.04)	(0.04)	(0.04)	(0.05)
Employment status dummies (baseline: Employed for wages)				
Self-employed	.143***	.138***	.138***	.144***
	(0.02)	(0.02)	(0.02)	(0.03)
Out of work for more than 1 year	-.671***	-.67***	-.67***	-.69***
	(0.04)	(0.05)	(0.05)	(0.06)
Out of work for less than 1 year	-.556***	-.533***	-.531***	-.487***
	(0.03)	(0.04)	(0.04)	(0.05)
Homemaker	.126***	.119***	.114***	.097***
	(0.02)	(0.02)	(0.02)	(0.03)
Student	.159***	.153***	.148***	.142***
	(0.04)	(0.04)	(0.04)	(0.05)
Retired	.236***	.227***	.214***	.219***
	(0.02)	(0.02)	(0.02)	(0.03)
Unable to work	-.543***	-.54***	-.558***	-.601***
	(0.03)	(0.03)	(0.03)	(0.04)

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] A win by a team expected to lose by at least 9 points given the pre-game betting spread.
[2] The post-game window is a period of three days after the last game was played.
[3] Equal to 1 if there are children living in the household with the respondent.

(Continued on the next page)

Table 1.D.3: Placebo Regression: Ordered Logit Coefficients (continued)
Dependent Variable: Life Satisfaction

Education dummies (baseline: High school graduate)				
Never attended school or only kindergarten	-.323**	-.278*	-.273*	-.465*
	(0.15)	(0.16)	(0.16)	(0.25)
Grades 1 - 8 (Elementary)	-.103***	-.126***	-.154***	-.173***
	(0.03)	(0.03)	(0.03)	(0.04)
Grades 9 - 11 (Some high school)	-.04*	-.033	-.055**	-.078**
	(0.02)	(0.02)	(0.02)	(0.03)
College 1 to 3 years (Some college or technical school)	.046***	.045***	.045***	.037*
	(0.01)	(0.01)	(0.01)	(0.02)
College 4 years or more (College graduate)	.2***	.198***	.199***	.203***
	(0.01)	(0.02)	(0.01)	(0.02)
Income dummies (baseline: $35,000 to under $50,000)				
Annual household income under $10,000	-.45***	-.46***	-.479***	-.436***
	(0.03)	(0.03)	(0.03)	(0.05)
Annual household income $10,000 to under $15,000	-.392***	-.389***	-.404***	-.389***
	(0.03)	(0.03)	(0.03)	(0.04)
Annual household income $15,000 to under $20,000	-.321***	-.327***	-.343***	-.361***
	(0.02)	(0.02)	(0.02)	(0.03)
Annual household income $20,000 to under $25,000	-.276***	-.276***	-.282***	-.263***
	(0.02)	(0.02)	(0.02)	(0.03)
Annual household income $25,000 to under $35,000	-.137***	-.143***	-.147***	-.127***
	(0.02)	(0.02)	(0.02)	(0.03)
Annual household income $50,000 to under $75,000	.16***	.16***	.165***	.18***
	(0.02)	(0.02)	(0.02)	(0.02)
Annual household income over $75,000	.451***	.442***	.451***	.45***
	(0.02)	(0.02)	(0.02)	(0.03)
Income info missing	-.021	-.029	-.044**	-.054*
	(0.02)	(0.02)	(0.02)	(0.03)
Health proxies (smoking baseline: Never smoked)				
Physically exercising[4]	.346***	.345***	.356***	.348***
	(0.01)	(0.01)	(0.01)	(0.02)
Limited in activity[5]	-.709***	-.713***	-.713***	-.693***
	(0.01)	(0.01)	(0.01)	(0.02)
Smoking every day	-.348***	-.376***	-.378***	-.369***
	(0.02)	(0.02)	(0.02)	(0.02)
Smoking some days	-.253***	-.296***	-.298***	-.243***
	(0.02)	(0.03)	(0.02)	(0.04)
dum_smokM	.032***			
	(0.01)			
Quit smoking		-.033***	-.029**	-.024
		(0.01)	(0.01)	(0.02)
cut1	-4.72***	-4.92***	-5.12***	-5.02***
	(0.06)	(0.29)	(0.29)	(0.76)
cut2	-2.99***	-3.19***	-3.39***	-3.28***
	(0.06)	(0.29)	(0.29)	(0.77)
cut3	.345***	.143	-.046	.059
	(0.06)	(0.29)	(0.29)	(0.77)
Weekly fixed effects	No	Yes	Yes	Yes
Team-State fixed effects	No	No	Yes	Yes
Observations	191,350	171,236	171,236	82,847
Games Included	All	All	All	Home

Standard errors adjusted for clusters at the county level in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[4] Equal to 1 if a person participated in physical exercise outside work in past 30 days.
[5] Equal to 1 if activities were limited due to physical, mental, or emotional problems.

Chapter 2

Criminals on the Field: A Study of College Football

Radek Janhuba and Kristýna Čechová[1]

Abstract

Economists have found mixed evidence on what happens when the number of police increases. On one hand, more law enforcers means a higher probability of detecting a crime, which is known as the monitoring effect. On the other hand, criminals incorporate the increase into their decision-making process and thus may commit fewer crimes, constituting the deterrence effect. This study analyzes the effects of an increase in the number of on-field college football officials, taking players as potential criminals and officials as law enforcers. Analyzing a novel play-by-play dataset from two seasons of college football, we report evidence of the monitoring effect being present in the overall dataset. This effect is mainly driven by offensive penalties

[1]Kristýna Čechová is affiliated with the Institute of Economic Studies, Faculty of Social Sciences, Charles University. Earlier versions of this study were published as Janhuba, R. & Cechova, K. (2017) "Criminals on the Field: A Study of College Football", *CERGE-EI Working Paper Series No. 610*, and as Janhuba, R. & Cechova, K. (2017) "Criminals on the Field: A Study of College Football", *IES Working Paper 13/2017*. We thank Randall Filer, Jan Hanousek, Stepan Jurajda, participants in a CERGE-EI Brownbag seminar, participants in the ESEA 2017 Paderborn conference, and two anonymous referees for helpful comments and suggestions. Janhuba acknowledges the institutional support RVO 67985998 from the Czech Academy of Sciences. All remaining errors are our own.

which are called in the area of jurisdiction of the added official. Decomposition of the effect provides evidence of the presence of the deterrence effect in cases of penalties with severe punishment or those committed by teams with moderate to high ability, suggesting that teams are able to strategically adapt their behavior following the addition of an official.

2.1 Introduction

What is the effect of increasing the number of police on crime rates? Based on the economic model of crime established by Becker (1968), the decision to engage in criminal activities depends on the expected utility of committing a crime. Specifically, potential criminals make their decision based on possible benefits, costs (punishment), probability of conviction, and considering their individual specific characteristics such as education.

An increase in the number of police can increase the probability of being caught and therefore convicted. If this increase is unobserved by potential criminals, it leads to an increase in reported crime rates, constituting the monitoring effect. However, potential criminals will likely observe an increase in the number of police. They will therefore incorporate it into their decision making process, change their behavior, and decrease the number of crimes committed (as their expected utility has decreased). This is the deterrence effect. As the monitoring and deterrence effects have opposite directions, the total effect of increasing the number of police on reported crime rates can be either positive or negative depending on the magnitude of each effect.

Our study looks at sports as an environment in which players are potential criminals and officials take the role of law enforcers.[2] Examining a novel play-by-play dataset, we evaluate the effects of increasing the number of officials from seven to eight in the 2014 and 2015 National Collegiate Athletic Association (NCAA)

[2]For the purpose of keeping the terminology clear, we abstain from using the term *referee* for a person observing the game and policing the rules. Instead, the term *official* is used. This is because in our context there are seven or eight officials on the field, and the one in charge of the whole officiating crew is called the *Referee*. Throughout the study, we identify this official using the term *Referee* (with a capital R).

football seasons.[3] To the best of our knowledge, this is the first study to examine this policy change on a nation-wide dataset.

In the sports context,[4] the economic model of crime established by Becker (1968) has been examined by several studies modeling fouls committed by players and the number of officials. McCormick and Tollison (1984) show that adding a third official in college basketball led to a decrease in the number of penalties called. Although Hutchinson and Yates (2007) discovered that McCormick and Tollison's results are erroneous due to a coding mistake, the corrected results still present evidence in favor of the existence of the deterrence effect (McCormick and Tollison, 2007).

Levitt (2002) and Heckelman and Yates (2003) analyze an experiment in the National Hockey League in which, during the 1999-2000 season, games were observed by either one or two referees.[5] Both papers find that the number of penalties increased and thus show that the monitoring effect was stronger than the deterrence effect (if there was any deterrence effect at all). Levitt (2002) argues that the change in the probability of detection was too small to result in an observable deterrence effect. Heckelman and Yates (2003) conclude that breaking rules in sports may not be well thought out but rather impulsive.

The sports policy evaluations closest to ours were carried out by Kitchens (2014) and Kitchens et al. (2017). Kitchens (2014) analyzes a natural experiment in the National Football League, which moved the position of the official known as the Umpire from behind the defense to behind the offense, keeping the number of officials fixed at seven.[6] Their results reveal that, after the change in the spatial distribution of officials, the number of penalties called on offense increased by 14% while the number of penalties called on defense decreased by 17%.

Interestingly, the eighth official added in college games, the Center Judge, was added to the same place as the new NFL Umpire position, while the college Umpire

[3]Note that throughout this study, the word *football* refers specifically to American football.

[4]Models of actual criminal behavior have been examined in several studies, many of which, however, suffer from endogeneity. General studies on crime are not reviewed in this study, which is focused on testing the deterrence effect in a sports environment. For a thorough review of the economics of crime, see Paternoster (2010) or Chalfin and McCrary (2014).

[5]In hockey terminology, the term *official* is not widely used. Instead, the game is supervised by referees and linesmen.

[6]The experiment may be viewed as natural because the primary reason for moving the Umpire's position was his safety, which is unrelated to the number of penalties called.

stayed in his original location. Thus, our results may be viewed as complementary to the results of Kitchens (2014) because the policy change we analyze added an official to a specific location, while Kitchens's analysis combined this intervention with the removal of an official from a different location.

Kitchens et al. (2017) study the policy change we analyze (see Section 2.2.2 for a description). On the dataset from the 2012 and 2013 seasons and studying games played by teams that were in the Big 12 Conference, they find evidence of the monitoring effect and limited support for the existence of the deterrence effect. Our study examines the 2014 and 2015 seasons and extends the sample to include all FBS football games.[7]

The contributions of our study are threefold. First, we add to the empirical literature on the strength and existence of the monitoring and deterrence effects. By identifying specific types of penalties, we can isolate the two effects. Second, our results indicate that there is a strategic interaction of teams following the policy change. Third, this is the first study to examine the policy change in question on a nationwide dataset.[8]

Our results indicate the presence of the monitoring effect in the overall dataset. This result is strengthened by performing a decomposition based on the area of the officials' jurisdiction. We also find evidence of the existence of the deterrence effect in two scenarios. First, we find an indication of the deterrence effect in cases of penalties carrying severe punishments. This may be explained by teams adapting their behavior as a response to the policy change. Second, we find limited evidence of the deterrence effect present in cases of non-severe penalties when only teams with moderately high (albeit not the highest) ability are considered. This may indicate that only teams at a relatively high playing level are able to strategically change their behavior.

The remainder of this study is structured as follows. Section 2.2 provides a brief introduction to the rules of football and to the intervention. Section 2.3 describes the dataset. Section 2.4 considers the methodology used. Section 2.5 presents the

[7]FBS is the highest level of college football played in the United States.

[8]Thus, our study may be seen as complementary to Kitchens et al. (2017), who examine the same policy change, but only for the Big 12 conference and for the period before our sample starts.

results. Section 2.6 concludes.

2.2 Football Specifics and Intervention Details

This section first introduces the sport of (American) football and its specifics that are important for this study. It then describes the intervention and discusses its implications. Readers familiar with the game of football may prefer to skip the next section and proceed directly to Section 2.2.2.

2.2.1 The Game of Football

Football is a collective sport played with 11 players on two teams on a rectangular field divided by lines into a grid. The last zone on each end of the field is known as the end zone.

The game is conducted in short consecutive plays usually lasting only seconds. After each play, the ball is placed either on the spot where it was at the end of the play, or the spot where the previous play started, depending on the outcome of the play. The team which initiates the ball into play is called the *offense*, and its objective is to get the ball into the opponent's end zone in order to score. The opposing team protects its end zone in order to keep the offense from scoring, and is called the *defense*.

When a team is awarded the ball, it has four opportunities (*downs*) to move the ball at least 10 yards closer to the opponent's end zone. If the offense succeeds, the down count resets and the offense again has a first down and 10 yards to go.[9] If the offense fails to achieve the first down during the four attempts, the ball is turned over to the defense at the spot where the fourth attempt ended.[10] The defense is then awarded a 1st down and hence becomes the offense, and vice-versa.

[9] For example if the situation is labeled a *2nd* (down) & *5*, the team has a second down and must advance the ball at least 5 yards to get the first down. If the team advances the ball 3 yards only, the next down will be labeled the *3rd* & *2*. If the team advances the ball 10 yards, they will get *1st* & *10* at the spot where the play ends.

[10] Teams will very rarely attempt to get the first down on a fourth down. Instead, they usually elect to try a field goal (see the next paragraph) or *punt* the ball, in which case they kick it towards the opponent's end zone so that the other team will need to gain a larger distance to score.

The goal of the game is to score more points than the opposing team. Kicking the ball through the uprights of the "Y" shaped goal results in a field goal worth 3 points. A touchdown worth 7 points is scored by moving the ball into the opponent's end zone by either carrying it there or catching it there.[11] Finally, a safety worth 2 points is awarded to the opposing team if a team is stopped in its own end zone (this occurs very rarely).

The games are governed by seven, or recently eight, officials.[12] These officials observe the game and if they see rule violations,[13] they throw a yellow flag to indicate a penalty. After the play ends, they confer and the Referee (the official responsible for the whole crew) then informs the teams and spectators of their decision. The usual form of penalty is a loss of 5, 10, or 15 yards, according to the severity of the foul.[14] The penalty is then assessed against the fouling team and the down is repeated.[15]

2.2.2 Change in the Number of Officials

We analyze a policy change in which the number of officials overseeing football games increased from seven to eight. The intervention was implemented gradually over three seasons. In the 2013 season, eight-member officiating crews oversaw exclusively games governed by the Big 12 conference. In the 2014 season, an additional three conferences[16] adopted the same rule change, while in 2015 it was applied to the whole FBS. The gradual introduction enables us to study the intervention as a

[11] Technically, the scoring team receives 6 points for a touchdown. Afterwards, it attempts one more play (so called "extra point" or "try") for which it can receive one point for kicking a field goal, an outcome that happens almost all the time, or two points if it scores another touchdown. An unsuccessful try for either a field goal or a touchdown means that the team receives 6 points for the touchdown.

[12] Note that this holds only for the highest level of college football games. Lower level college and professional games are governed by seven officials.

[13] Although the basic rules of the game are quite simple, the specifics of play are governed by a complex set of rules (e.g., the 2016 official NCAA football rule book contains 218 pages of text).

[14] In specific circumstances, the penalty can shorten in distance by taking the form of half the distance to the goal line or by placing the ball on the spot where the foul was committed.

[15] For example, if the offense commits a holding foul on 2nd & 10 which results in a gain of 15 yards, the gain is canceled and the next down will be 2nd & 20. In some specific cases, the penalty can also include a loss of a down for offensive penalties or an automatic 1st down for defensive infractions.

[16] Specifically, these were the Atlantic Coast Conference, the Big 10 Conference and the American Athletic Conference.

natural experiment.

The first policy change of this type since 1983 was adopted as a response to an increase in the speed of the game in the previous years, and related issues.[17] The officials began to have difficulty preparing for the next play quickly enough to assure proper observation of the game. Player safety and potential gaps in coverage were also widely discussed topics.[18]

Generally, officials have divided areas of coverage, meaning that each official has a specific area to observe and detailed instructions on what types of fouls to watch for in particular. Gaps in this coverage meant that specific fouls were missed due to the seven officials not being able to observe all the actions taking place on the field. Note that while no official is strictly restricted from calling fouls that occur outside his area of jurisdiction, the officials are specifically trained to observe their area exclusively, and are actively criticized when they take actions outside of their jurisdiction. Note also that the officials work in crews that remain together for the entire season and they are thus generally well aware of what their colleagues would be doing in each specific situation that may arise during the football game.

A graphical illustration of the change in the composition of officials is depicted in Figure 2.2.1.[19] The added official has been labeled a Center Judge and is positioned in the offensive backfield, behind the offense and to the opposite side from the quarterback than the Referee. Thus, his area of jurisdiction mainly includes fouls in the area of the offensive line (broadly defined) and defensive fouls against the quarterback.

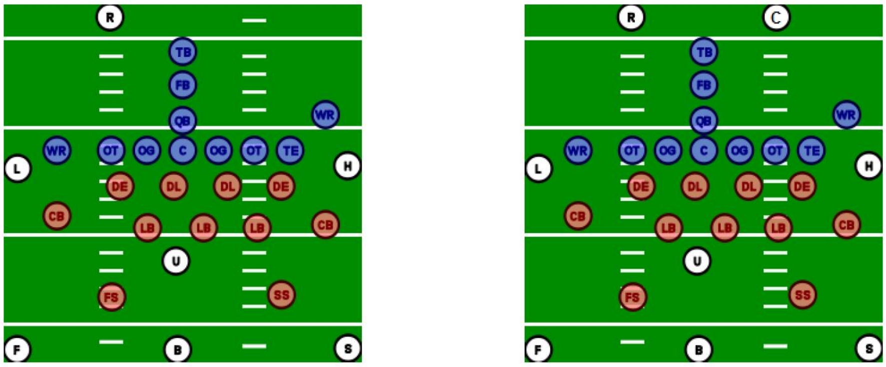
---

[17]The increase in pace was associated with the implementation of the 40-second clock. The 40-second clock rule introduced in 2008 sets an interval between the end of a play and the beginning of a new one at no more than 40 seconds. Previously, the clock was only 25 seconds but counting began only after the officials made the ball ready for play. The aim of the rule was to increase the pace of games (Source: http://bleacherreport.com/articles/35981-2008-rule-changes-what-every-fan-needs-to-know).

[18]For examples, see: http://www.cbssports.com/college-football/news/big-12-adds-eighth-official-just-to-keep-up-with-up-tempo-offenses/ or http://www.cbssports.com/college-football/news/sec-to-experiment-with-8-football-officials-but-whats-right-number/

[19] The schematics of the American football officiating positions has been downloaded from Wikimedia Commons under the Creative Commons Attribution-Share Alike 3.0 Unported license, and was subsequently modified to illustrate the policy change. Author: Derivative work by Zzyzx11 based on the original image by UserB. Detailed information: https://commons.wikimedia.org/wiki/File:American_football_officials_positions.svg

Figure 2.2.1: Schematics of the Policy Change



(a) Before **(b)** After

Source: Wikimedia Commons (see footnote 19 for details)

Although the main reason for implementing the policy change was unrelated to penalty-related behavior, there is still a potential threat that conferences which voluntarily adopted the policy change in 2014 differed from those that waited until 2015. Therefore, we performed balancing tests for penalty-related statistics in the season before the intervention took place.

Specifically, we examined the overall aggregated team levels of penalty-related measures from the 2013 season, and analyzed whether their distribution differs for conferences that initiated the eighth official in 2014. We also checked if the two conference groups differed in the speed of play before the intervention took place. The results of these balancing tests are presented in Table 2.2.1. Both the t-tests and Kolmogorov-Smirnov tests for equality of distributions suggest that the distribution of balancing characteristics did not differ between conferences that did and did not implement the policy change in 2014. Hence, we conclude that the intervention may be viewed as exogenous to penalty characteristics.

To keep the decision as clean as possible, our reported results exclude the Big 12 and independent teams.[20] Keeping the Big 12 in the dataset would mean that one of the balancing groups would include data influenced by the "trial run" of the intervention. Nevertheless, results from balancing tests including the Big 12 teams in the treatment group are qualitatively the same.[21]

[20]In 2013 there were six so-called independent teams. These teams are not governed by any conference.

[21]These results are available upon request.

Table 2.2.1: Balancing Tests

	Control[1]		Treatment[1]		t-test		K-S test	
	Mean	SD	Mean	SD	t-stat	p-val	D stat	p-val
Penalties per game	5.56	1.16	5.69	1.24	-0.52	0.60	0.10	0.96
Penalties per play	0.04	0.01	0.04	0.01	-0.88	0.38	0.13	0.81
Penalty yards per game	48.03	10.73	48.44	11.44	-0.18	0.85	0.10	0.96
Penalty yards per play	0.34	0.07	0.34	0.08	-0.56	0.58	0.11	0.92
Plays per game	143.33	8.98	141.05	7.65	1.30	0.19	0.15	0.67
Plays in season	1817.27	167.03	1797.14	114.48	0.65	0.52	0.20	0.27

Tests performed on aggregate data in the 2013 season. Calculation excludes independent and Big 12 teams.
[1] Treatment includes teams that adopted the intervention in 2014. Control includes teams that did not.
Source: Authors' calculation; Data from `http://www.cfbstats.com/2013/team/index.html`

2.3 Data

The data on football games have been downloaded from the NCAAsavant.com website.[22] The data include play-by-play information for NCAA football games in the 2014 and 2015 seasons.[23] Note that the dataset was mainly created as the base of an interactive website and unfortunately does not cover games from the 2013 season (when all games were governed by seven officials), which causes methodological issues (see below).[24]

The dataset includes basic variables about each play, such as which team is on the offense, the type of play, the result of the play, and a detailed text description of the play. This text includes information about penalties called during the play. Therefore, we can identify the penalty type, team, player, and whether the specific penalty was called on the offense or defense.

The aggregated seasonal statistics for each team were obtained from the website of SportSource Analytics.[25] The information from this source contains the aggregates for the number of offensive plays, offensive yards, number of penalties, and penalty yards in the seasons from 2008 to 2015. All of these variables are available

[22] `http://ncaasavant.com`

[23] More precisely, the data present a subsample of football games in each season. In the 2014 season, the missing games seem to be random. In the 2015 season, the dataset covers the first seven weeks of the season.

[24] We attempted to contact the owner of the website to obtain access to codes used to compile the dataset, which would allow us to obtain the same data for the 2013 season (as well as missing games from 2014 and 2015). The owner did not respond to our inquiries.

[25] `http://www.cfbstats.com/`

for each team and for its opponents.

The data on officiating assignments were downloaded from the collegiate athletics websites of all 128 universities that were part of the FBS in the 2014 and 2015 seasons. Note that while play-by-play statistics are generally available from sports news websites for our sample period, these servers usually do not include data on which officiating crew supervised the particular game. This information is available in the official game statistics, which the home team is required to collect and upload to the NCAA. The teams then release this official report on their athletics websites. Note that as the officials work in crews that are constant through the entire season, we only record the name of the Referee to identify which officiating group oversaw the specific game.

After matching the three data sources the main dataset includes 148,097 plays from 1,011 games. To simplify the analysis, we decided to restrict the dataset to basic plays from scrimmage (rushes and passes).[26]

The descriptive statistics for play-by-play data are presented in Table 2.3.1. The first two rows show the proportions of run and pass plays. The last three rows show the unconditional probability of a penalty occurring, followed by the probability of penalties for offensive holding and roughing the passer, which constitute the two specific types of fouls we are particularly interested in (see the next section for explanation).[27]

Table 2.3.1: Descriptive Statistics

	Mean	S.D.	Min	Max	N
Running play	0.5107	0.4999	0	1	148,097
Passing play	0.4893	0.4999	0	1	148,097
Any penalty	0.0463	0.2100	0	1	148,097
Offensive holding	0.0125	0.1109	0	1	148,097
Roughing the passer	0.0029	0.0536	0	1	72,462

Source: Authors' calculation

[26]Thus, we eliminate plays involving kicks. Although these are undoubtedly an important part of a football game, the behavior of players during kick plays is substantially different and their inclusion would introduce noise into the analysis.

[27]Note that the number of observations for roughing the passer penalties is approximately half of the number for other variables. This is due to the fact that this type of penalty can only appear in passing plays, while the other types can appear in runs as well as in passes.

2.4 Methodology

Using the play-by-play information, we examine the probability of a specific penalty being called within every play. The basic model takes the form

$$y_{ighvr} = \lambda_1 \left[\text{eight}_g\right] [2014] + \lambda_2 \left[\text{eight}_g\right] [2015] + \beta X_{ig} + \theta_h + \theta_v + \theta_r + \varepsilon_{ighvr} \quad (2.1)$$

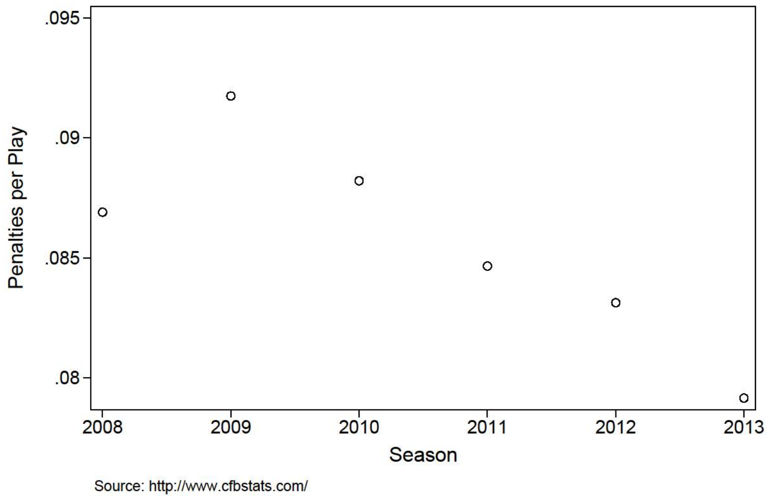
where the subscript $ighvr$ can be read as "play i in game g of home team h and visiting team v under the supervision of Referee r". The dependent variable y is an indicator equal to one if the specific type of penalty was called within the play. The variables in brackets mark indicators equal to one when the condition described by the inside of the bracket is specified. Specifically, `eight` is an indicator equal to one if the game was supervised by eight officials, and `2014` with `2015` are indicators equal to one if the game was played in a particular season. X is a vector of football specific variables for each play, namely, distance to first down, field position, and indicator variables for down, quarter, and whether the play was a run or pass. Finally, θ_h, θ_v, and θ_r are fixed effects for the home team, the visiting team, and the officiating crew represented by the Referee.

The particular regression methodology has been selected in order to perform two types of comparisons. First, the coefficient λ_1 captures the within-season variation of adding an extra official and thus can capture the immediate adjustment to the new number of officials. Second, the coefficient λ_2 captures the between-season variation and measures the effect of introducing the policy for all games in the 2015 season.

Note that while a such methodology would be ideal, it is not plausible to estimate the effect using the standard difference-in-differences framework, as the dataset we possess does not have information on games played in the 2013 season.

Note also that as the policy change influenced all observations in the second year of the sample, it is difficult to disentangle the effect of the intervention and a potential time trend in the dependent variable. More specifically, if there is a time trend in the dependent variable, one should look at the coefficient and then deduct the time trend from the estimated value of the effect. Figure 2.4.1 shows the values of the number of penalties divided by the number of offensive plays and its evolution in years 2008 through 2013.[28] The figure reveals that there is a negative time trend in the number of penalties per play.[29] Therefore, as the regression design inherently assumes that there is a zero time trend, the empirical results will likely tend to underestimate the true effect rather than to overestimate it. In other words, as there is likely a negative time trend in the dependent variable, a potentially positive regression coefficient should arguably be viewed more credibly than it would if it had a negative value.

Figure 2.4.1: Pre-Treatment Time Trend in Penalties



In combination with the specific characteristics of several types of football fouls, the intervention allows us to shed light on the difference between the monitoring and deterrence effects. Specifically, these are offensive holdings and roughing the passer penalties, both of which occur predominantly in the area of the new official's jurisdiction. The following paragraphs explain why these two types of penalties can be used for a deeper analysis.

First, holding seems to be the type of foul which is most likely to be influenced by

[28]Due to data limitations, the unconditional probability of penalties is measured in a different setting and is therefore not comparable to the values in Table 2.3.1. This is because the play-by-play data for previous seasons is not available and we can therefore only make inferences based on the total number of penalties called on each team including dead-ball fouls such as false starts and/or penalties that are called during kick plays.

[29]The existence of the negative trend is supported by ordinary least squares results of average penalty rates on time.

the policy change. Specifically, before the change, the Umpire and the Referee were assigned responsibility for fouls occurring in the area of the offensive line. The basic assignment decomposition was that the Umpire observed fouls committed by the three interior linemen, while the Referee observed fouls by both exterior linemen. Clearly, this was often an impossible job, and consequently the Referee observed only one of the two potential suspects. The introduction of the third observer in the area means that all relevant players can be observed at all times.

Moreover, while it is impossible to prove that there will be no deterrence effect at all, it is also arguably likely that it would be negligible in cases of offensive holding. This is because offensive holding occurs in practice when the defensive player outplays his offensive rival, who resorts to illegal holding so as not to allow his opponent to continue to move in the direction of the ball carrier. In fact, coaches often instruct players, especially in cases of passing plays, to hold the opponent rather than allow him to continue towards the quarterback, as a holding penalty punishes the team by 10 yards but avoids the potential tragedy of injury to the key player.[30]

Therefore, we find it reasonable to suspect that the number of offensive holdings called would have risen following the introduction of the extra official. In terms of the economic model of crime, while there is a higher probability of being caught, the benefits of committing the crime outweigh the potential penalty.

The second type of penalty we are interested is *roughing the passer*, which occurs when the defender hits the quarterback after he has released the ball. The reason is that the second backfield official sees the passer from a second angle and can thus help to cover this safety-related foul, and therefore the officials are less likely to miss them (constituting a higher probability of detection).

Additionally, such penalties carry an automatic first down for the offense and a risk of the responsible player being disqualified from playing in the remainder of the game in cases of serious misconduct. Therefore, due to the severity of the punishment, roughing the passer penalties should arguably be associated with a

[30]The quarterback is the most important player on the team and injury to him may have catastrophic consequences for the team in question.

stronger deterrence effect. Thus, in terms of the economic model of crime, roughing the passer fouls are crimes with a high punishment.

2.5 Results

2.5.1 All Penalties

The results of the linear probability model regressions for all penalties are presented in Table 2.5.1.[31] The results indicate that although the number of penalties increased

Table 2.5.1: Linear Probability Model: All Penalties

	(1)	(2)	(3)	(4)
Eight-men crew in 2014	-0.0009	-0.0007	-0.0006	-0.0006
	(0.0018)	(0.0017)	(0.0019)	(0.0021)
Eight-men crew in 2015	0.0026	0.0029*	0.0035**	0.0043**
	(0.0017)	(0.0017)	(0.0017)	(0.0020)
Yards to 1st down		0.0007***	0.0006***	0.0006***
		(0.0002)	(0.0002)	(0.0002)
Field position		-0.0001***	-0.0001***	-0.0001***
		(0.0000)	(0.0000)	(0.0000)
Passing play		0.0276***	0.0273***	0.0273***
		(0.0012)	(0.0012)	(0.0012)
Constant	0.0455***	0.0229***	0.0396***	0.0467***
	(0.0015)	(0.0024)	(0.0131)	(0.0128)
Down and Quarter	No	Yes	Yes	Yes
Teams	No	No	Yes	Yes
Referee	No	No	No	Yes
N	148,097	147,192	147,192	147,192

Standard errors adjusted for 101 clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Estimation of the model.

in the 2015 season following the policy change, this increase is not visible in the games supervised by eight officials during the 2014 season. Hence, the result may not be solely attributable to the presence of the new official. As there has been

[31] We present linear probability model regressions due to the direct interpretation of regression coefficients as marginal effects and their lower computational time required. Robustness of the estimation method is discussed in section 2.5.5.

no other major rule change between the two seasons, one possible interpretation of this result is that the officials may have known that the policy change poses a new issue for the teams and they subsequently "went easy" on the players for the 2014 season. A second possible interpretation is linked with the fact that increasing the number of officials necessarily meant that the newly added official did not have prior experience at the same level.

2.5.2 Areas of Officiating Coverage

Due to the specific spatial allocation of football officials and their areas of coverage, results on all penalties combined may be imprecise as they include fouls which occurred in areas not observed by the new official. Specifically, as the new official was added into the area behind the offense, he would typically be expected to call more penalties on the offense and fewer on the defense.[32] Therefore, we redefined the dependent variable into two separate indicators equal to one when the penalty was called on offense or on defense, and repeated the estimation.

Moreover, in order to analyze the situation in the greatest possible detail, we further devoted our attention to two types of penalties which should arguably be most influenced by the extra official. These are offensive holding and roughing the passer penalties. As explained in Section 2.4, analysis of these specific penalties should provide insights into the existence of the deterrence effect.

The results are presented in Table 2.5.2. The first two columns reveal that, as expected, the effect can be mainly attributed to increases in offensive penalties. Specifically, while the effect is statistically insignificant in the 2014 season, the probability of a penalty called within a play supervised by seven officials is 0.0215. Hence, the effect of 0.0045 called under the supervision of eight officials in the 2015 season represents an increase of 21.1 percent.

Due to the spatial allocation of the officials, the regression reported in the second column may be viewed as a placebo test. We can see that, as the defensive penalties remain the same following adoption of the eighth official, the placebo test indicates

[32] An exception is roughing the passer penalties, which are explored below.

Table 2.5.2: Linear Probability Model: Area of Coverage

	Offensive Penalties (1)	Defensive Penalties (2)	Offensive Holding (3)	Offensive PI[1] (4)	Roughing the Passer (5)
Eight-men crew in 2014	0.0021	-0.0027	0.0000	-0.0011	0.0014**
	(0.0014)	(0.0016)	(0.0011)	(0.0008)	(0.0007)
Eight-men crew in 2015	0.0045***	-0.0002	0.0020**	-0.0006	0.0006
	(0.0014)	(0.0014)	(0.0009)	(0.0006)	(0.0006)
Yards to 1st down	0.0006***	0.0000	0.0005***	-0.0002***	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Field position	-0.0000	-0.0001***	0.0000	-0.0000*	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	0.0007	0.0267***	-0.0063***		
	(0.0009)	(0.0009)	(0.0007)		
Constant	0.0109	0.0357***	-0.0004	0.0146***	0.0030
	(0.0074)	(0.0086)	(0.0043)	(0.0038)	(0.0049)
N	147,192	147,192	147,192	71,964	71,964

The dependent variable is specified by the column heading.
All columns include the full set of fixed effects for down, quarter, teams, and Referee.
Standard errors adjusted for 101 clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] PI stands for "Pass Interference".
Source: Estimation of the model.

that the increase in penalties is indeed driven by fouls committed on the offense.[33]

The offensive holding results, in which the deterrence effect is expected to be very small, are presented in the third column. We can see that the results for offensive holding penalties are qualitatively similar to the overall results for all penalties, as there is an effect of more penalties being called, but only after all of the conferences adopted the eighth official in the 2015 season.[34] Specifically, while the effect is statistically insignificant in the 2014 season, the probability of an offensive holding being called within a play supervised by seven officials is 0.0121, while the effect of 0.002 represents an increase of 16.8 percent called under the supervision of eight officials in the second season. This result is qualitatively similar to the result of

[33]Following Kitchens (2014), we also performed a placebo test for defensive holding on runs. As the coefficient on both variables in question was insignificant at the 5% level, the qualitative implications of this alternative placebo specification remain the same.

[34]We also attempted restricting the sample to offensive holdings called during passing plays only. The results are qualitatively identical.

Kitchens (2014), who finds an increase of 14% following the relocation of the umpire from behind the defense to behind the offense.

The fourth column provides another placebo test by looking at offensive pass interference penalties, which are arguably the only type of offensive fouls that should not be even theoretically influenced by the eighth official.[35] The fact that the coefficients are insignificant in both periods validates the finding from the regression in the third column.

The results reported in the fifth column seemingly suggest that the number of roughing the passer penalties increased following the policy change, but only in the 2014 season when some games were still supervised by seven officials. However, as discussed in more detail below in section 2.5.3, the coefficient on the variable in the 2014 season is statistically significant only because the 2015 season games are included in the regression and thus influence the regression benchmark. Hence, the positive value of the coefficient in the 2014 season actually picks up the deterrence effect occurring between the two seasons. This may be explained by the fact that, in connection with roughing the passer penalties, it is difficult to change one's behavior when only selected games are supervised by eight officials, while it is more possible to establish a behavioral change between the two seasons.

To sum up, the results in this section present evidence of an overall monitoring effect and suggest the existence of the deterrence effect in the case of crimes with the most serious punishment.

2.5.3 Experience with the Policy

To better understand the strategic interactions with regard to the policy change, we now look at regressions estimated on weekly subsamples. Due to the policy taking universal effect in 2015, the regressions estimated in this section only look at 2014. Moreover, the results in this section exclude games including the Big 12 teams from the sample, as these teams had already played under the supervision

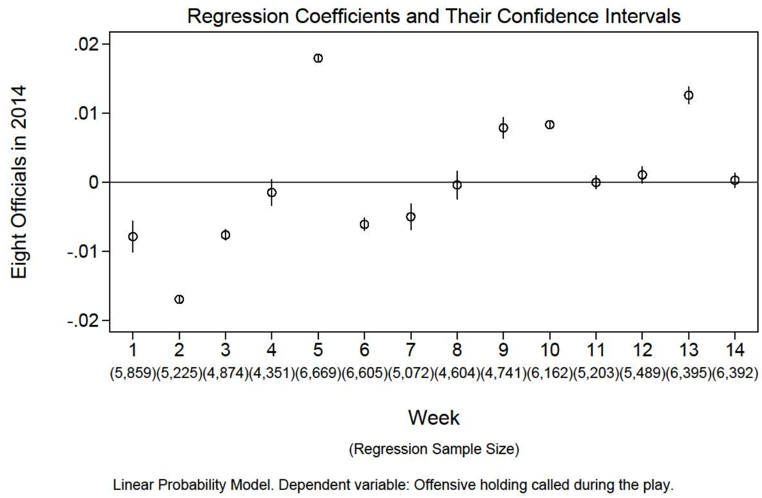
[35] This is because when a pass is thrown, all officials except for the Referee and Center Judge look towards the area where the ball will land. The two remaining officials observe the quarterback, looking out for the roughing the passer penalty. Thus, in the case of offensive pass interference, the presence of the new official carries zero spillover effect.

of eight officials in the 2013 season.[36] Due to the results being extensive in terms of the space needed to show regression outputs, we present the coefficients on the treatment variable graphically. Full results are available upon request.

Offensive Holding

The effects on offensive holding are presented in Figure 2.5.1. The figure reveals an interesting pattern during the 2014 season, in which the original regression coefficient was statistically insignificant. The weekly decomposition reveals that, with the exception of one week, there is a clear upward trend in the overall number of reported offensive holdings throughout the 2014 season.

Figure 2.5.1: Experience and Offensive Holding



One way to explain this trend is that players are being coached to hold their opponent if they cannot block him legally. With the addition of the extra official in the offensive backfield, players likely started to fear a penalty and thus decreased the number of holding fouls. However, as the season progressed, they were coached not to adjust their behavior and gradually reverted to the overall stable level of fouls committed. Thus, throughout the season, the number of offensive holding fouls increased and the monitoring effect prevailed towards the end of the season.

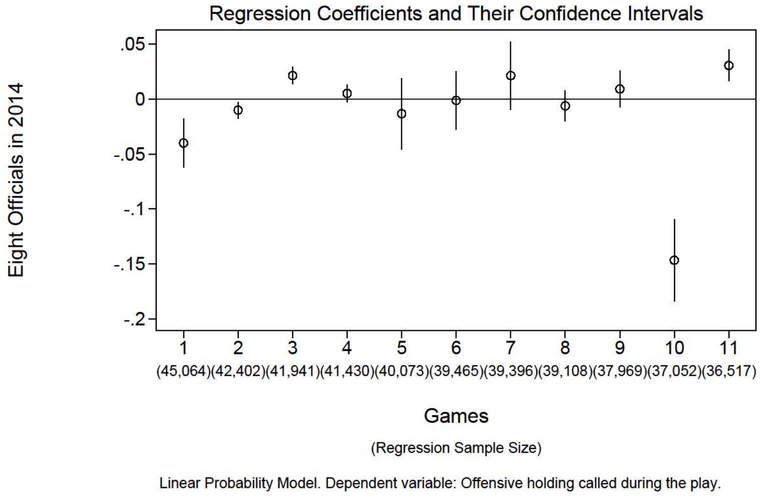
In order to isolate whether the effect lies in players' ability to adjust their behavior, we also looked at the effects of the policy broken down by the number of games

[36]Including the Big 12 teams does not substantially alter the results.

that the crew has officiated up to and including the game in question. For each level of officiating experience, the control group includes all the games supervised by seven officials in the weeks included in the treatment group.

The results of this breakdown are shown in Figure 2.5.2. The figure reveals that there is no clear trend in the number of fouls called based on the experience of the officiating crew in question. Therefore, we conclude that the upward trend in offensive holding penalties is likely caused by the strategic interaction of teams in response to the new policy.

Figure 2.5.2: Officiating Experience and Offensive Holding

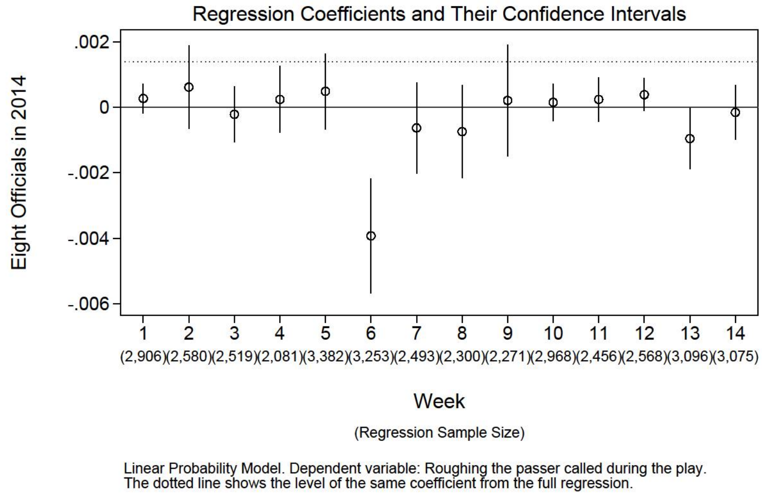


Roughing the Passer

Results on roughing the passer penalties are presented in Figure 2.5.3. We can see that there is no clear in-season trend in these cases. More importantly, notice that all of the weekly coefficients are lower in magnitude than the overall effect of the policy in the 2014 season of 0.014. This serves as evidence of the overall coefficient in the roughing the passer regression reported in column 5 of Table 2.5.2 actually picking up the deterrence effect of the 2015 season, rather than the monitoring effect of the 2014 season. This is caused by the overall benchmark being influenced by the games from the 2015 season as well.[37]

[37]Indeed, when one estimates the roughing the passer regression on the data from the 2014 season only, the coefficient on the eight-men crew is statistically insignificant with a p-value of 0.35.

Figure 2.5.3: Experience and Roughing the Passer



2.5.4 Role of Team Quality

We now extend the analysis to let the effect of the intervention on the two specific types of penalties differ based on team quality. This is motivated by the possibility that high- and low-skilled teams differ in their game strategies and ability to adjust their behavior in response to the policy change.

In order to distinguish between the offensive ability of the teams, we took the total yards gained by each team's offense in the previous season and ranked the teams according to their performance. Analogically, we took the total offensive yards gained by the opposing teams to evaluate defensive abilities. We then defined the best teams as the top 25 teams in each category. This selection is motivated by the fact that college sports usually rank the best 25 teams overall. The robustness of the number of teams belonging in the top category is discussed in Section 2.5.5.

The results are shown in Table 2.5.3. The first two columns suggest that the number of offensive holdings already decreased with the addition of the eighth official in 2014, but only for the teams with a high quality offense. This result indicates the presence of the deterrence effect within teams with high offensive ability.

The third and fourth columns report the roughing the passer analysis broken down to whether the defensive team belongs to the top 25 teams. The effect is insignificant for the teams with the highest defensive quality. This can either mean

Table 2.5.3: Breakdown by Team Quality (LPM)

	Offensive Holding		Roughing the Passer	
	Top 25 Offense (1)	Other Offense (2)	Top 25 Defense (3)	Other Defense (4)
Eight-men crew in 2014	-0.0078***	0.0007	0.0023	0.0016**
	(0.0026)	(0.0011)	(0.0021)	(0.0007)
Eight-men crew in 2015	-0.0033	0.0028***	0.0008	0.0010*
	(0.0025)	(0.0011)	(0.0018)	(0.0006)
Yards to 1st down	0.0007***	0.0005***	0.0001	0.0001
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Field position	0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	-0.0084***	-0.0057***		
	(0.0014)	(0.0007)		
Constant	-0.0445***	0.0035	0.0050	-0.0002
	(0.0106)	(0.0051)	(0.0082)	(0.0049)
N	30,282	116,910	15,220	56,744

Columns are separated by the rankings based on own (opponents') yards gained (allowed) in the previous season.
All columns include the full set of fixed effects for down, quarter, teams, and Referee.
Standard errors adjusted for 84 (Top 25) or 101 clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Estimation of the model.

that there was no effect for these teams, or that (similarly to offensive holding) the high-quality teams alter their behavior. However, given that the overall roughing the passer rates are the same across the two categories of teams, the difference between the two coefficients could arguably be caused by teams' ability to change their strategic behavior after the policy change. In other words, if there is a deterrence effect, it is likely present for the relatively high-skilled teams.

2.5.5 Robustness Checks

Match Pair Fixed Effects

In this section we perform an alternative specification of the regressions. Instead of including a set of fixed effects for the home team and a second set of fixed effects for the visiting team, we keep only those games in which teams played against each other in both years and include a fixed effect for a match-pair combination, ignoring which team played at home and which was on the road.

The results are reported in Table 2.5.4. Interestingly, all coefficients almost double in magnitude. Moreover, the results for all penalties and offensive holding are more precisely estimated due to the benchmark being more specifically set, using match-pair fixed effects.[38]

In the case of roughing the passer coefficients, its value is larger in magnitude but less precisely estimated.[39] Nevertheless, the reported value is qualitatively consistent with other results.

Note also that due to the data being available only for a subsample of games, the reduction of the sample size is substantial. The fact that the results hold even after this decrease in the number of observations further confirms the validity of our results.

Number of Top Teams Considered

Although it is customary to rank the top 25 teams in college sports, the choice of splitting the sample to the best 25 teams remains arbitrary. The sensitivity of the coefficient on eight officials in the second season based on the number of top offensive teams is depicted in Figure 2.5.4. Full regression results are presented in Table 2.A.2 in Appendix 2.A.

[38]Note that the term "more precisely estimated" is meant in connection with the absolute value of the coefficient. In other words, it does not correspond to a tighter confidence interval, but rather to a result with a higher statistical significance.

[39]Because it was the 2014 coefficient that was significant for roughing the passer regressions, the alternative specification with match-pair fixed effects brings less precision into the estimation of this coefficient. This is because due to the structure of the competition, no two teams played against each other twice in one season.

Table 2.5.4: Match Pair Fixed Effects

	All penalties		Offensive Holding		Roughing the Passer	
	Team	Match-pair	Team	Match-pair	Team	Match-pair
	(1)	(2)	(3)	(4)	(5)	(6)
Eight-men crew in 2014	-0.0006	0.0036	0.0000	0.0014	0.0014**	0.0021*
	(0.0021)	(0.0039)	(0.0011)	(0.0015)	(0.0007)	(0.0011)
Eight-men crew in 2015	0.0043**	0.0082***	0.0020**	0.0044***	0.0006	0.0014
	(0.0020)	(0.0029)	(0.0009)	(0.0012)	(0.0006)	(0.0009)
Yards to 1st down	0.0006***	0.0002	0.0005***	0.0006***	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Field position	-0.0001***	-0.0001**	0.0000	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	0.0273***	0.0283***	-0.0063***	-0.0069***		
	(0.0012)	(0.0019)	(0.0007)	(0.0011)		
Constant	0.0192	0.0288*	-0.0062	0.0072	0.0083**	0.0034
	(0.0154)	(0.0154)	(0.0058)	(0.0063)	(0.0033)	(0.0043)
N	147,192	45,211	147,192	45,211	71,964	22,394

Columns headed by "Team" come from the baseline specification. Columns headed by "Match-pair" only include teams that played each other twice during the sample period.
All columns include the full set of fixed effects for down, quarter, teams, and Referee.
Standard errors adjusted for 81 (Match-pair) or 101 clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Estimation of the model.

We can see that the result found for the top 25 teams holds if we relax the condition for the 25 teams towards a higher number, but, somewhat surprisingly, it does not hold if we constrain the estimation to only the few teams with the strongest offense. Even more surprisingly, analysis of the top 5, 10, and 15 teams suggests exactly the opposite scenario for these teams. A possible explanation for this is that the best teams are so good that they do not need to adjust their behavior in the sense that they are not afraid of a penalty being called. Alternatively, it may be that the top 15 teams are the most skilled and aggressive, and hence the addition of the extra official means that there is a stronger monitoring effect.

The sensitivity of the coefficient for the effect during the 2014 season on the roughing the passer penalties is examined in Figure 2.5.5. The full regression results are reported in Table 2.A.3 in Appendix 2.A.

Figure 2.5.4: Offensive Holding and Number of Top Teams

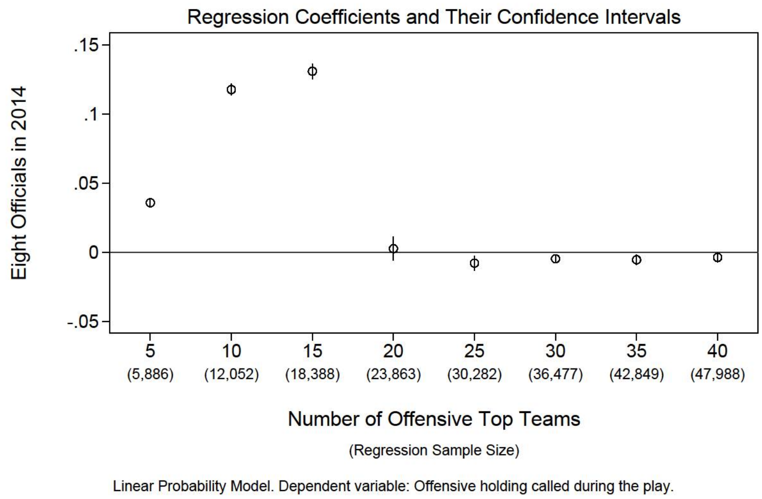
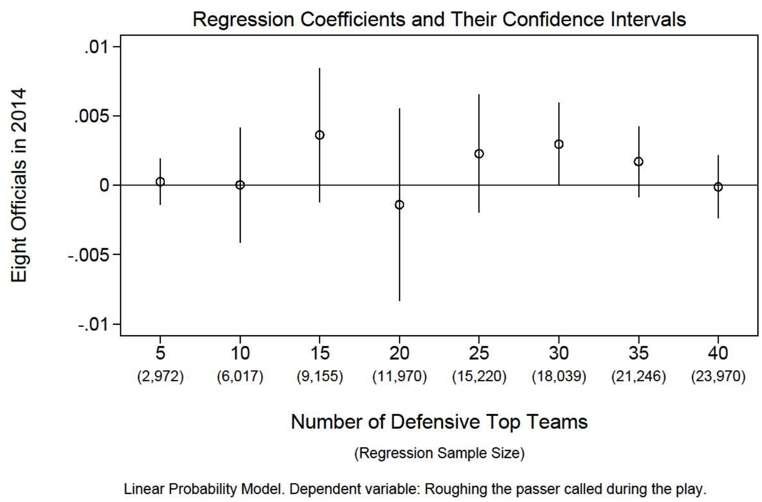


Figure 2.5.5: Roughing the Passer and Number of Top Teams



We can see that in the roughing the passer case the interpretation of a possible presence of the deterrence effect holds regardless of how many of the top defensive teams are considered. Interestingly, the very top teams show an increase in the number of roughing the passer penalties in the second season (see Table 2.A.3).

Logit Specification

Due to the direct interpretation of regression coefficients and the less computational time required, all results reported in the main text come from a linear probability model. In order to check for robustness to alternative functional form, the analysis was re-estimated using a logit specification.

Note that, unfortunately, some of the specifications are not estimable with a logit specification while keeping the full set of fixed effects. This is because there are either very few observations where the dependent variable is equal to one (such as in the case of the roughing the passer penalties) and/or the sample size is not large enough to include the full sets of fixed effects into the analysis.[40]

In order to overcome this issue, we re-estimated these models without the Referee fixed effects. The problem, however, persisted in the specific case of defensive team quality regressions, in which there was a combination of small sample size and very small proportion of penalties in the sample. Therefore, in these regressions we kept the Referee fixed effects in the model but excluded team fixed effects. In all cases in which the logit estimation was performed using a different set of fixed effects, we re-estimated the same regression using the linear probability model as well.

The comparison of the marginal effects can be found in Appendix 2.B. All results are qualitatively identical to their linear probability model counterparts. Therefore, we conclude that there is no severe functional form specification issue in the analysis.

2.6 Conclusion

This study evaluates the effects of the policy change of increasing the number of collegiate football officials from seven to eight in the highest level of NCAA football. Comparing our results with the previous literature, this is the first study to find evidence of both the monitoring and deterrence effect on a nationwide dataset.

Analyzing a play-by-play dataset from two seasons of college football games, we report evidence of a monitoring effect being present in the overall dataset. Moreover, analysis of offensive holding and roughing the passer penalties, which constitute misconduct that is especially likely to be observed by the added official, also suggests that there is a monitoring effect present.

We also report evidence of the deterrence effect being present in two scenarios. First, we find an indication of the deterrence effect in the roughing the passer penal-

[40]Including the full set of fixed effects causes the model likelihood to be flat. Hence, convergence is not achieved when estimating the full form of the model.

ties during the second observed season. This is likely caused by between-season changes in team behavior. Second, we find limited evidence of the deterrence effect being present in both types of penalties when only teams with moderate to high ability are considered. This indicates that teams with high (albeit not the highest) skills are able to strategically interact based on the policy change.

The results are robust to alternative specification of fixed effects, functional form of the estimation, and the number of teams considered in the relatively high-skilled group.

Appendix

2.A Appendix 2.A: Number of Top Teams

Table 2.A.1 reports the top teams in the offensive and defensive rankings for each of the two seasons analyzed.

Tables 2.A.2 and 2.A.3 show full regression results from regressions on subsamples on the top teams discussed in Section 2.5.5. Specifically, the coefficients in the first row of Table 2.A.2 are depicted in Figure 2.5.4. Figure 2.5.5 shows coefficients in the second row of Table 2.A.3.

The sample restriction was performed based on the total number of offensive yards gained in the previous season, with higher numbers of a given team's yards equating to a better offensive ranking. In defensive rankings, the team that allowed the lowest overall number of opponents' offensive yards was ranked the highest.

Table 2.A.1: Overview of Top Teams

Rank	Offensive Top Teams		Defensive Top Teams	
	2014	2015	2014	2015
1	Baylor	Oregon	Louisville	Clemson
2	Oregon	Marshall	Michigan State	Penn State
3	Northern Illinois	Ohio State	Virginia Tech	Stanford
4	Florida State	Baylor	Alabama	Michigan
5	Ohio State	West. Kentucky	Florida	UCF
6	Fresno State	East Carolina	Florida State	Florida
7	Auburn	TCU	Iowa	Louisville
8	Marshall	Boise State	Florida Atlantic	Michigan State
9	Texas A&M	Alabama	Wisconsin	Wisconsin
10	Missouri	Mississippi State	West. Kentucky	LSU
11	Texas Tech	Georgia Tech	Cincinnati	Temple
12	Clemson	Wisconsin	South Florida	App. State
13	Colorado State	Michigan State	TCU	Syracuse
14	Washington	West Virginia	Wake Forest	Arkansas
15	Bowling Green	Arizona	LSU	Boston College
16	BYU	Toledo	Memphis	Virginia
17	Arizona State	Auburn	Bowling Green	Mississippi
18	Georgia	Colorado State	UTSA	Miami (Florida)
19	Wisconsin	Wash. State	North Texas	San Jose State
20	Ball State	Arkansas State	Mississippi State	San Diego State
21	Boise State	Florida State	South Carolina	Florida Intl.
22	Mississippi	Northern Illinois	Oklahoma	Buffalo
23	Cincinnati	UCLA	Tulane	Georgia
24	Indiana	Bowling Green	Penn State	Wake Forest
25	East Carolina	Texas Tech	South Alabama	North Texas
26	Oregon State	Oklahoma	Vanderbilt	UTSA
27	Louisville	BYU	Connecticut	TCU
28	Duke	Cincinnati	Clemson	Akron
29	Arizona	Georgia	Utah State	Houston
30	San Jose State	USC	Kansas State	Virginia Tech
31	Alabama	California	Baylor	Iowa
32	LSU	Texas A&M	Georgia Tech	LA Monroe
33	South Carolina	Nebraska	USC	Texas
34	Oklahoma State	GA Southern	UCF	Connecticut
35	UCLA	Notre Dame	Notre Dame	Memphis
36	Utah State	South Carolina	Texas State	Alabama
37	UCF	Arizona State	Akron	Northwestern
38	Stanford	West. Michigan	Utah	Central Michigan
39	Rice	Fresno State	Pittsburgh	Tulane
40	Wyoming	Pittsburgh	Syracuse	GA Southern

Offensive teams ranked by the most total offensive yards in the specific season.
Defensive teams ranked by the least total opponents' offensive yards in the specific season.
Data from `http://www.cfbstats.com/`

Table 2.A.2: Offensive Holding Regression: Number of Top Offensive Teams Considered

	Top 5	Top 10	Top 15	Top 20	Top 25	Top 30	Top 35	Top 40
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eight-men crew in 2014	0.0358***	0.1176***	0.1308***	0.0027	-0.0078***	-0.0046***	-0.0053***	-0.0037**
	(0.0015)	(0.0021)	(0.0028)	(0.0044)	(0.0026)	(0.0016)	(0.0018)	(0.0018)
Eight-men crew in 2015	0.0477***	0.1933***	0.2134***	0.0104***	-0.0033	-0.0001	-0.0019	-0.0015
	(0.0015)	(0.0030)	(0.0042)	(0.0034)	(0.0025)	(0.0017)	(0.0017)	(0.0017)
Yards to 1st down	0.0006	0.0008**	0.0010***	0.0008***	0.0007***	0.0007***	0.0006***	0.0007***
	(0.0006)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Field position	0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	-0.0081**	-0.0086***	-0.0075***	-0.0081***	-0.0084***	-0.0091***	-0.0082***	-0.0077***
	(0.0033)	(0.0022)	(0.0018)	(0.0016)	(0.0014)	(0.0013)	(0.0012)	(0.0012)
Constant	-0.0296***	-0.2217***	0.1170***	-0.0244***	-0.0445***	-0.0325***	-0.0306***	-0.0281***
	(0.0057)	(0.0040)	(0.0061)	(0.0077)	(0.0106)	(0.0106)	(0.0107)	(0.0089)
N	5,886	12,052	18,388	23,863	30,282	36,477	42,849	47,988

All columns include the full set of fixed effects for down, quarter, teams, and Referee.

Standard errors adjusted for 48 (Top 5) to 92 (Top 40) clusters by the Referee in parentheses.

Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Estimation of the model.

Table 2.A.3: Roughing the Passer Regression: Number of Top Defensive Teams Considered

	Top 5	Top 10	Top 15	Top 20	Top 25	Top 30	Top 35	Top 40
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eight-men crew in 2014	0.0003	0.0000	0.0036	-0.0014	0.0023	0.0030*	0.0017	-0.0001
	(0.0008)	(0.0021)	(0.0024)	(0.0035)	(0.0021)	(0.0015)	(0.0013)	(0.0011)
Eight-men crew in 2015	0.0007	0.0129***	0.0183***	-0.0011	0.0008	0.0007	-0.0001	-0.0014
	(0.0009)	(0.0032)	(0.0038)	(0.0026)	(0.0018)	(0.0016)	(0.0013)	(0.0011)
Yards to 1st down	0.0003	0.0003	0.0003	0.0001	0.0001	0.0000	0.0001	0.0000
	(0.0005)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Field position	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-0.0049	0.0016	-0.0144**	0.0153	0.0050	0.0090	-0.0002	0.0035
	(0.0062)	(0.0044)	(0.0059)	(0.0106)	(0.0082)	(0.0071)	(0.0071)	(0.0066)
N	2,972	6,017	9,155	11,970	15,220	18,039	21,246	23,970

All columns include the full set of fixed effects for down, quarter, teams, and Referee.

Standard errors adjusted for 48 (Top 5) to 92 (Top 40) clusters by the Referee in parentheses.

Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Estimation of the model.

2.B Appendix 2.B: Logit Results

Tables in this section present regression results based on alternative functional form using logit estimation. Otherwise, specification of all tables is identical. With the exception of Table 2.2.5, the particular sub-number of all tables corresponds to the sub-numbers of tables in Section 2.5.[41]

Note that in columns marked by subscripts ^{R} or ^{T} in the heading of the columns, the regression does not include Referee or team fixed effects and is therefore not directly comparable to the appropriate regression in the main text. This is because some of the specifications were not estimable using a logit while keeping the fixed effects due to either a small number of observations with the dependent variable equal to one or insufficient sample size causing the likelihood function to become flat.

As discussed in Section 2.5.5, in order to establish the validity of comparisons in the case described in the previous paragraph, we decided to re-estimate these models without Referee or team fixed effects using both the linear probability model and logit specifications. Results of these regressions are shown in Table 2.2.5.

[41] Thus, for example, Table 2.B.2 corresponds to Table 2.5.2.

Table 2.B.1: Marginal Effects from Logit Model: All Penalties

	(1)	(2)	(3)	(4)
Eight-men crew in 2014	-0.0009	-0.0008	-0.0007	-0.0002
	(0.0018)	(0.0017)	(0.0020)	(0.0021)
Eight-men crew in 2015	0.0025	0.0028*	0.0035**	0.0046**
	(0.0017)	(0.0017)	(0.0017)	(0.0020)
Yards to 1st down		0.0006***	0.0005***	0.0005***
		(0.0001)	(0.0001)	(0.0001)
Field position		-0.0001***	-0.0001***	-0.0001***
		(0.0000)	(0.0000)	(0.0000)
Passing play		0.0285***	0.0281***	0.0282***
		(0.0013)	(0.0013)	(0.0013)
N	148,097	147,192	147,192	147,192

Standard errors adjusted for 101 clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Estimation of the model.

Table 2.B.2: Marginal Effects from Logit Model: Area of Coverage

	Offensive Penalties (1)	Defensive Penalties (2)	Offensive Holding (3)	Offensive PI[1R] (4)	Roughing the Passer[R] (5)
Eight-men crew in 2014	0.0024*	-0.0033*	0.0002	-0.0014	0.0009
	(0.0014)	(0.0017)	(0.0012)	(0.0012)	(0.0013)
Eight-men crew in 2015	0.0047***	-0.0007	0.0022**	-0.0009	-0.0001
	(0.0013)	(0.0015)	(0.0009)	(0.0009)	(0.0012)
Yards to 1st down	0.0005***	-0.0000	0.0005***	-0.0003***	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Field position	-0.0000	-0.0001***	0.0000	-0.0000**	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	0.0008	0.0301***	-0.0064***		
	(0.0009)	(0.0011)	(0.0007)		
N	147,192	146,639	145,402	44,484	43,918

The dependent variable is specified by the column heading.
All columns include the full set of fixed effects for down, quarter, teams, and Referee.
Standard errors adjusted for 101 clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1] PI stands for "Pass Interference".
[R] Robust regression estimated without Referee fixed effects.
Source: Estimation of the model.

Table 2.B.3: Breakdown by Team Quality (Logit MEs)

	Offensive Holding		Roughing the Passer	
	Top 25 Offense[R] (1)	Other Offense (2)	Top 25 Defense[T] (3)	Other Defense[T] (4)
Eight-men crew in 2014	-0.0089***	0.0012	0.0017	0.0001
	(0.0028)	(0.0013)	(0.0040)	(0.0008)
Eight-men crew in 2015	-0.0021	0.0032***	-0.0024	0.0004
	(0.0027)	(0.0011)	(0.0029)	(0.0008)
Yards to 1st down	0.0006***	0.0004***	0.0001	0.0001
	(0.0002)	(0.0001)	(0.0003)	(0.0001)
Field position	0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	-0.0100***	-0.0058***		
	(0.0017)	(0.0007)		
N	25,630	115,353	7,479	45,269

Columns are separated by the rankings based on own (opponents') yards gained (allowed) in the previous season.
All columns include the full set of fixed effects for down, quarter, teams, and Referee.
Standard errors adjusted for clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[R] Robust regression estimated without Referee fixed effects.
[T] Robust regression estimated without team fixed effects.
Source: Estimation of the model.

Table 2.B.4: Match Pair Fixed Effects (Logit MEs)

	All penalties		Offensive Holding		Roughing the Passer	
	Team	Match-pair	Team	Match-pair	Team^{R}	Match-pair^{R}
	(1)	(2)	(3)	(4)	(5)	(6)
Eight-men crew in 2014	-0.0002	0.0024	0.0002	0.0011	0.0009	-0.0000
	(0.0021)	(0.0041)	(0.0012)	(0.0017)	(0.0013)	(0.0027)
Eight-men crew in 2015	0.0046**	0.0079**	0.0022**	0.0046***	-0.0001	-0.0002
	(0.0020)	(0.0031)	(0.0009)	(0.0015)	(0.0012)	(0.0020)
Yards to 1st down	0.0005***	0.0001	0.0005***	0.0005***	0.0001	0.0003
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0003)
Field position	-0.0001***	-0.0001**	0.0000	0.0000	-0.0000	-0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play	0.0282***	0.0294***	-0.0064***	-0.0071***		
	(0.0013)	(0.0021)	(0.0007)	(0.0012)		
N	147,192	45,410	145,402	42,603	43,918	8,003

Columns headed by "Team" come from baseline specification. Columns headed by "Match-pair" only include teams that played each other twice during the sample period.
All columns include the full set of fixed effects for down, quarter, teams, and Referee.
Standard errors adjusted for clusters by the Referee in parentheses.
Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
^{R} Robust regression estimated without Referee fixed effects.
Source: Estimation of the model.

Table 2.2.5: Regressions with Different Sets of Fixed Effects: Comparison of LPM and Logit Marginal Effects

	(1) to (8): Without Referee Fixed Effects								(9) to (12): Without Team Fixed Effects			
	Offensive PI[1]		Rough. the Passer		Rough. the Passer (Match-pair FE)		Offensive Holding (Top 25 Offense)		Rough. the Passer (Top 25 Defense)		Rough. the Passer (Other Defense)	
	LPM	Logit	LPM	Logit	LPM	Logit	LPM	Logit	LPM	Logit	LPM	Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Eight-men crew in 2014	-0.0009	-0.0014	0.0007	0.0009	-0.0001	-0.0000	-0.0007	-0.0089***	0.0008	0.0017	0.0001	0.0001
	(0.0008)	(0.0012)	(0.0007)	(0.0013)	(0.0010)	(0.0027)	(0.0023)	(0.0028)	(0.0017)	(0.0040)	(0.0006)	(0.0008)
Eight-men crew in 2015	-0.0007	-0.0009	-0.0000	-0.0001	-0.0001	-0.0002	0.0011	-0.0021	-0.0011	-0.0024	0.0003	0.0004
	(0.0005)	(0.0009)	(0.0005)	(0.0012)	(0.0008)	(0.0020)	(0.0021)	(0.0027)	(0.0011)	(0.0029)	(0.0006)	(0.0008)
Yards to 1st down	-0.0002***	-0.0003***	0.0001	0.0001	0.0001	0.0003	0.0008***	0.0006***	0.0001	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0001)	(0.0001)
Field position	-0.0000*	-0.0000**	-0.0000	-0.0000	-0.0000	-0.0001	0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Passing play							-0.0081***	-0.0100***				
							(0.0014)	(0.0017)				
N	71,964	44,484	71,964	43,918	22,511	8,003	30,282	25,630	15,220	7,479	56,744	45,269

Standard errors adjusted for clusters by the Referee in parentheses.

Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Estimation of the model.

Chapter 3

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

Radek Janhuba and Jakub Mikulka[1]

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.

[1]We 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. This study was supported with institutional support RVO 67985998 from the Czech Academy of Sciences. All remaining errors are our own.

3.1 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.[2]

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 et al., 2006).[3] In our setting, the stake size can be

[2]See also Feess et al. (2016) who show that empirical specification controlling for the average stake size leads to more precise estimates of bettors' risk preferences.

[3]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 and Tamarkin, 1998)

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.[4]

Moreover, in experiments focused on the presence of hot-hand and gambler's fallacies in betting behavior, Matarazzo et al. (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 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.

---

or probability misperceptions (Snowberg and 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).

[4]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 and Thaler (2003). See also Moskowitz (2015).

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 3.2 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 3.2.1). We describe the dataset in Section 3.3 and the empirical methodology in Section 3.4. Section 3.5 discusses the empirical results, and Section 3.6 analyzes their robustness. Section 3.7 concludes.

3.2 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 3.2.2.

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).

3.2.1 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 3.2.1, 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 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.

Table 3.2.1: 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
			$E(X) = 90$	$E(X) = 81$

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

$$E_{single} = \sum_{k=1}^{N} \pi_k w_k = \sum_{k=1}^{N} \frac{1}{x_k} \frac{M}{N} x_k = M, \tag{3.1}$$

while the expected payoff from placing an accumulator bet is

$$E_{accum} = \pi_w w = \left(\prod_{k=1}^{N} \frac{1}{x_k} \right) M \left(\prod_{k=1}^{N} x_k \right) = M. \tag{3.2}$$

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

$$E_{single} = \sum_{k=1}^{N} \pi_k w_k = \sum_{k=1}^{N} \frac{1 - \delta}{x_k} \frac{M}{N} x_k = (1 - \delta)\, M, \tag{3.3}$$

$$E_{accum} = \pi_w w = \left(\prod_{k=1}^{N} \frac{1-\delta}{x_k} \right) M \left(\prod_{k=1}^{N} x_k \right) = (1-\delta)^N M. \tag{3.4}$$

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.

3.2.2 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 and 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 and 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.[5] 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.

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

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. 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.

3.3 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.[6]

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.[7] Arguably, the market is thus

[6]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.

[7]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

comparable to markets in areas where gambling regulation restricts betting to only 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 3.3.1. 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 3.3.1: 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 3.3.2 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

the market in general.

overround being compounded in favor of the bookmaker (see equation 3.4 in Section 3.2.1).

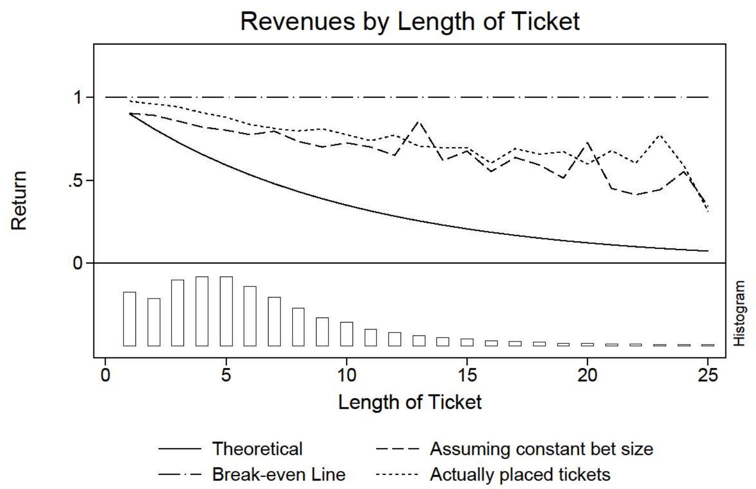
Table 3.3.2: Descriptive Statistics - Ticket Level

	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 3.3.1. 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.

Figure 3.3.1: Length and Profitability of 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 motivation for testing our research question, i.e. the effect of the stake size on the probability of success and net revenues.

3.4 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

$$y_{itk} = \beta_0 + \beta X_{itk} + \gamma C_{itk} + \alpha_i + \theta_t + \varepsilon_{itk} \tag{3.5}$$

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.[8] The combination of time and user fixed effects controls for the possible presence of behavioral biases that were previously found in sports betting markets

[8]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.

(e.g. the home bias documented by Braun and 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 large-scale limited dependent variable models with fixed effects.

3.5 Results

3.5.1 Baseline Analysis

The results of the baseline analysis are shown in Table 3.5.1. 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.

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

Table 3.5.1: 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.

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.[9]

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.

[9]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.

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 3.2.1, 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 more profitable to bet due to the favorite longshot bias as discussed in Section 3.4. 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 3.5.3.

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 3.6.2.

The results are presented in Table 3.5.2. In Panel A, the main result of the stake being positively correlated with the success of individual bets is valid for all of the bettors except for those with the lowest skills.[10] Even though the sample is divided into groups of the same number of bettors, bettors with the lowest skills also place

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

Table 3.5.2: 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.

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.

3.5.3 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 3.5.3. 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.

Table 3.5.3: 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

To see whether the effect of stake size on the success rate and net revenues differ 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 3.5.4.

Table 3.5.4: 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.

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 3.5.3, 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 3.5.4, 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.

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.

3.5.4 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 3.5.5 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 3.5.6 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.

Table 3.5.5: 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 3.5.6: 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.

3.6 Robustness Checks

3.6.1 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 3.6.1. 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 3.5.1.

Table 3.6.1: 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.

3.6.2 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 3.6.2 and 3.6.3. 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 3.6.2: 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.

Table 3.6.3: 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.

The results for net revenue reveal a slightly different pattern. Specifically, there 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.[11] Hence, the indirect measure of skills via betting activity does not seem to be a particularly reliable one.

3.6.3 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

[11] 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.

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 3.6.4. 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 3.6.4. 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 placed a bet at least every month.

The results are presented in Table 3.6.5. 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 3.6.5 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.

Table 3.6.4: 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 3.6.5: 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.

3.7 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.

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 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.

Bibliography

Agarwal, S., Duchin, R., and Sosyura, D. (2013). In the Mood for a Loan: The Causal Effect of Sentiment on Credit Origination.

Ahlfeldt, G. M. and Kavetsos, G. (2014). Form or function?: the effect of new sports stadia on property prices in London. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(1):169–190.

Ahlfeldt, G. M. and Maennig, W. (2010). Impact of sports arenas on land values: evidence from Berlin. *The Annals of Regional Science*, 44(2):205–227.

Allmers, S. and Maennig, W. (2009). Economic impacts of the FIFA Soccer World Cups in France 1998, Germany 2006, and outlook for South Africa 2010. *Eastern Economic Journal*, 35(4):500–519.

Andrikogiannopoulou, A. and 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.

Ansolabehere, S., Rodden, J., and Snyder, James M., J. (2006). Purple America. *Journal of Economic Perspectives*, 20(2):97–118.

Baade, R. A., Baumann, R. W., and Matheson, V. A. (2008). Assessing the Economic Impact of College Football Games on Local Economies. *Journal of Sports Economics*, 9(6):628–643.

Baade, R. A., Baumann, R. W., and Matheson, V. A. (2011). Big Men on Campus: Estimating the Economic Impact of College Sports on Local Economies. *Regional Studies*, 45(3):371–380.

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

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The Economic Dimensions of Crime*, pages 13–68. Springer.

Blanchflower, D. G. and Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, 66(8):1733–1749.

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

Card, D. and Dahl, G. B. (2011). Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Chalfin, A. and McCrary, J. (2014). Criminal deterrence: A review of the literature. *Journal of Economic Literature*. In press.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352.

Chen, D. L. (2016). This Morning's Breakfast, Last Night's Game: Detecting Extraneous Influences on Judging. *Social Science Research Network Working Paper Series*.

Coates, D. and Humphreys, B. R. (2008). Do Economists Reach a Conclusion on Subsidies for Sports Franchises, Stadiums, and Mega-Events? *Econ Journal Watch*, 5(3):294–315.

Deaton, A. and Stone, A. A. (2014). Evaluative and hedonic wellbeing among those with and without children at home. *Proceedings of the National Academy of Sciences*, 111(4):1328–1333.

Depetris-Chauvin, E., Durante, R., and Campante, F. R. (2018). Building Nations Through Shared Experiences: Evidence from African Football. Working Paper 24666, National Bureau of Economic Research.

Doerrenberg, P. and Siegloch, S. (2014). Is soccer good for you? The motivational impact of big sporting events on the unemployed. *Economics Letters*, 123(1):66 – 69.

Dolan, P., Peasgood, T., and White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, 29(1):94 – 122.

Drake, M. S., Gee, K. H., and Thornock, J. R. (2016). March Market Madness: The Impact of Value-Irrelevant Events on the Market Pricing of Earnings News. *Contemporary Accounting Research*, 33(1):172–203.

Edmans, A., García, D., and Norli, Ø. (2007). Sports Sentiment and Stock Returns. *The Journal of Finance*, 62(4):1967–1998.

Ely, J., Frankel, A., and Kamenica, E. (2015). Suspense and Surprise. *Journal of Political Economy*, 123(1):215–260.

Eren, O. and Mocan, N. (2018). Emotional Judges and Unlucky Juveniles. *American Economic Journal: Applied Economics*, 10(3):171–205.

Fair, R. C. and Oster, J. F. (2007). College Football Rankings and Market Efficiency. *Journal of Sports Economics*, 8(1):3–18.

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

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

Fernquist, R. M. (2000). An aggregate analysis of professional sports, suicide, and homicide rates: 30 U.s. metropolitan areas, 1971-1990. *Aggression and Violent Behavior*, 5(4):329 – 341.

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

Gainsbury, S. M. and 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. and 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 Siegel, D. S. and Vaughan-Williams, L., editors, *The Oxford Handbook of the Economics of Gambling*, pages 341–369. Oxford University Press.

Healy, A. J., Malhotra, N., and Mo, C. H. (2010). Irrelevant events affect voters' evaluations of government performance. *Proceedings of the National Academy of Sciences*, 107(29):12804–12809.

Heckelman, J. C. and Yates, A. J. (2003). And a hockey game broke out: Crime and punishment in the nhl. *Economic Inquiry*, 41(4):705–712.

Huang, H. and Humphreys, B. R. (2012). Sports participation and happiness: Evidence from US microdata. *Journal of Economic Psychology*, 33(4):776 – 793.

Humphreys, B. R. and Nowak, A. (2017). Professional sports facilities, teams and property values: Evidence from NBA team departures. *Regional Science and Urban Economics*, 66(C):39–51.

Hutchinson, K. and Yates, A. (2007). Crime on the court: A correction. *Journal of Political Economy*, 115(3):515–519.

Jones, M. V., Coffee, P., Sheffield, D., Yangüez, M., and Barker, J. B. (2012). Just a game? changes in english and spanish soccer fans' emotions in the 2010 world cup. *Psychology of Sport and Exercise*, 13(2):162 – 169.

Kavetsos, G. and Szymanski, S. (2010). National well-being and international sports events. *Journal of Economic Psychology*, 31(2):158–171.

Kerr, J. H., Wilson, G. V., Nakamura, I., and Sudo, Y. (2005). Emotional dynamics of soccer fans at winning and losing games. *Personality and Individual Differences*, 38(8):1855 – 1866.

Kitchens, C. (2014). Identifying changes in the spatial distribution of crime: Evidence from a referee experiment in the national football league. *Economic Inquiry*, 52(1):259–268.

Kitchens, C., Makofske, M. P., and Wang, L. (2017). Parallel Lives of Hard Hitting Criminals: Evidence from NCAA Football. Working paper, Unpublished.

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.

Koszegi, B. and Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.

Kugler, T., Kausel, E. E., and Kocher, M. G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(4):471–482.

Levitt, S. D. (2002). Testing the Economic Model of Crime:The National Hockey League's Two-Referee Experiment. *The B.E. Journal of Economic Analysis & Policy*, 1(1):1–21.

Lindo, J. M., Siminski, P., and Swensen, I. D. (2018). College Party Culture and Sexual Assault. *American Economic Journal: Applied Economics*, 10(1):236–65.

Matarazzo, O., Carpentieri, M., Greco, C., and 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*, pages 353–368. Springer.

McCormick, R. E. and Tollison, R. D. (1984). Crime on the Court. *Journal of Political Economy*, 92(2):223–35.

McCormick, R. E. and Tollison, R. D. (2007). Crime on the Court, Another Look: Reply to Hutchinson and Yates. *Journal of Political Economy*, 115(3):520–521.

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

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

Paternoster, R. (2010). How much do we really know about criminal deterrence? *The Journal of Criminal Law and Criminology*, pages 765–824.

Rees, D. I. and Schnepel, K. T. (2009). College Football Games and Crime. *Journal of Sports Economics*, 10(1):68–87.

Sauer, R. D. (1998). The economics of wagering markets. *Journal of Economic Literature*, 36(4):2021–2064.

Schwarz, N., Strack, F., Kommer, D., and Wagner, D. (1987). Soccer, rooms, and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific domains. *European Journal of Social Psychology*, 17(1):69 – 79.

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

Song, C., Boulier, B. L., and Stekler, H. O. (2007). The comparative accuracy of judgmental and model forecasts of American football games. *International Journal of Forecasting*, 23(3):405–413.

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

Süssmuth, B., Heyne, M., and Maennig, W. (2010). Induced Civic Pride and Integration. *Oxford Bulletin of Economics and Statistics*, 72(2):202–220.

Yechiam, E., Telpaz, A., and Hochman, G. (2014). The complaint bias in subjective evaluations of incentives. *Decision*, 1(2):147.

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