MDSAA

Master Degree Program in
Data Science and Advanced Analytics

IMPACT OF VIDEO ASSISTANT REFEREE ON PORTUGUESE FOOTBALL

Studying the effect on different game dynamics

Diogo Ferreira Gonçalves

Master Thesis

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

IMPACT OF VIDEO ASSISTANT REFEREE ON PORTUGUESE FOOTBALL

Studying the effect on different game dynamics

By
Diogo Ferreira Gonçalves

Master Thesis presented as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Business Analytics

Supervised by

Bruno Damásio, PhD, Universidade de Lisboa

November 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Diogo Gonçalves

Lisbon, 30 November 2023

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor Bruno Damásio, my supervisor, for his guidance and flexibility during the entire dissertation writing process. I would also like to extend my gratitude to Pedro Peres, who has always been available to provide clarifications on the data provided by the Federação Portuguesa de Futebol.

To my loving parents, brother and Mariana, my deepest appreciation for encouragement and motivation.

ABSTRACT

This study investigates the effects of the Video Assistant Referee on Portuguese top-tier football dynamics, including goal scoring, home advantage, playing time and competitiveness, employing a robust research approach encompassing descriptive analyses, non-parametric tests, Event Studies with multiple models, Generalized Linear Models and Regression Discontinuity Design. The key findings reveal that Video Assistant Referee has marginally increased average goals per game, primarily favouring away teams, while home advantage remains largely unaffected. Playing time has significantly increased post Video Assistant Referee implementation, with effective playing time showing a slightly increase. In terms of league competitiveness, Video Assistant Referee introduction hasn't substantially altered the winner margin on games. This study provides novel contributions, such as a comprehensive examination of home advantage using diverse metrics, a unique focus on effective playing time and one of the initial assessments of Video Assistant Referee impact on football competitiveness. A critical discovery suggests a potential link between increased playing time due to Video Assistant Referee and an elevated risk of player injuries, particularly within a context of rising match frequency. Additionally, the study raises concerns about the efficacy of extended playing time in enhancing game dynamics, hinting at possible diminishing returns in effective playing time. In summary, this research establishes a foundation for further exploration into the broader implications of Video Assistant Referee in football.

KEYWORDS

Video Assistant Referee; Football; Home Advantage; Non-Parametric Tests; Event Studies; Generalized Linear Models; Regression Discontinuity Design.

SUSTAINABLE DEVELOPMENT GOALS

INDEX

List of Figures

List of Tables

List of Abbreviations and Acronyms

CIES	International Centre for Sports Studies
ES	Event Study
FIFA	Fédération Internationale de Football Association
GLM	Generalized Linear Models
HA	Home Advantage
IFAB	International Football Association Board
RDD	Regression Discontinuity Design
VAR	Video Assistant Referee

1. Introduction

1.1. Football and Home Advantage

Football is widely known as the most popular sport in the world. It is estimated to be played by 200 million players across 200 countries. The sport's easiness of play, being able to be played anywhere, from Wimbledon to the neighbourhood streets and due to being free and accessible whenever there is a ball, make it global appealing and undisputed (Dvorak et al., 2004).

International football is regulated by FIFA, created in 1904 and composed by 211 national associations aggregated in 6 regional federations (FIFA, n.d.). Prior to FIFA, IFAB was created in 1886 and aims to develop and maintain the Laws of the Game. These two organizations aim to improve football's quality and fairness (IFAB, n.d.).

Yet, the question of fairness arises with the well-documented phenomenon of Home Advantage (HA). Right before the Referee's starting whistle, the home team have already more chances to win the game. There is consistent research demonstrating HA in almost all team sports (Courneya & Carron, 1992; Jamieson, 2010; Gómez et al., 2011; Jones, 2013; Ribeiro et al., 2016 and Pollard et al., 2017) and in subjective individual sports, despite to a lower extent (Ribeiro et al.,2016). Even the Olympic Games have evidence of HA (Balmer et al., 2001 and Franchini & Takito, 2016).

Football is one of the sports with the largest HA, as exhaustively studied by Pollard (1986), Pollard (2006), Jamieson (2010), Pollard and Gómez (2015) and Pollard et al (2017). It is considered, however, to be a decreasing phenomenon (Leite, 2017 and Peeters & van Ours, 2020).

Geography affects the extent of HA effect, varying across the globe (Pollard, 2006 and Pollard & Gómez, 2014). European leagues differ on the extent of HA, but there is evidence of it everywhere (Inan, 2018 and Marek & Vávra, 2020) and even in UEFA Champions League (Page & Page, 2007 and Lidor et al., 2010). Evidence however has not reached a consensus on whether lower levels have lower or higher home advantage (Sánchez et al., 2009, Leite & Pollard, 2018 and Marek & Vávra, 2020). Pollard and Gómez (2012) also found that women's league have lower home advantage effect than men's football.

1.2. Crowd Effect and Covid-19

The crowd presence is a longtime considered HA factor. In fact, there is solid literature backing the crowd effect on HA (Nevill & Holder, 1999; Goumas, 2014 and Ponzo & Scoppa, 2016) and its varying effect depending on crowd density (Anders & Rotthoff, 2014 and Inan, 2020).

Covid-19 pandemic provided natural experimental approaches on ghost games. Crowd support is shown to fasten the rate at which goals, cards and penalty kicks are given out (Magee & Wolaver, 2023). The probability of the home team missing a penalty increases when there is no crowd attendance, away teams are less likely to choke on a penalty (Ferraresi & Gucciardi, 2021) and HA decreased significantly, as shown by Tilp and Thaller (2020), Carlos Cueva (2020), CIES Football Observatory (2020), Ferraresi and Gucciardi (2020), Dilger and Vischer (2020), Endrich and Gesche (2020), Bryson et al. (2021), Reade et al. (2021), McCarrick et al. (2021), Leitner and Richlan (2021), Scoppa (2021), Morita and Araki (2022) and Leitner et al. (2023). This change was not observed for lower divisions (Fischer & Haucap, 2021). The reduction of HA was also seen on market expectation, reducing by one third the home team betting odds (Wunderlich et al.,2021).

There is however a few research on negative effect of crowd, which argues that home teams choke under pressure of larger audiences (Dohmen, 2008a; Boheim et al., 2018 and Harb-Wu & Krumer, 2019). Even further, home advantage was stated as existing regardless of a crowd being present (VAN DE VAN, 2011). However, none of these conclusions were confirmed on Covid-19 era.

Moreover, there are other factors believed to play a role on HA, as the existence of a running track in the stadium (BURAIMO, 2011 and Armatas & Pollard, 2014), distance travelled (Pollard, 2008; Pollard & Gómez, 2014 and Pollard et al., 2017), stadium altitude (Van Damme & Baert, 2019), pitch format (Barnett & Hilditch, 1993) and moving to new stadiums (Pollard, 2002). Also, playing at non frequent days reduces the HA for the underdog team (Goller & Krumer, 2020).

HA might be naturally intrinsic in the game. Even the total distance covered is bigger for players playing at home than when playing away (Castellano et al., 2011). Coaches, regardless of expertise, have higher expectations to win at home games (Staufenbiel et al., 2015).

1.3. Referees

Referees, on their side, are also prompted to be a source of home advantage and are found to be biased in several studies.

The research has collected evidence of crowd influence on referee decisions (Balmer et al., 2007; Petterson-Lidbom & Priks, 2010; Picazo-Tadeo et al.,2017 and Sapp et al.,2018). Referees tend to call more fouls and cards to the away team (Constantinou et al.,2014; Goumas, 2014; Picazo-Tadeo et al.,2017 and Sapp et al.,2018) and are biased in penalty kick situations (Erikstad & Johansen, 2020), which might be due to increased stress and anxiety (Balmer et al., 2007), once this bias differs between referees (Boyko et al., 2007 and Page & Page, 2010). Also, there is bias in favour home teams by shortening games where the home team is ahead (Garicano et al., 2001; Scoppa, 2007; Dohmen, 2008b; Riedl et al., 2015 and Lago-Peñas & Gómez-López, 2016).

HA is widely accepted as part of the game, but its origin is yet to be perfectly explained and far from reaching a consensus amongst the researchers. The most accepted conclusion is that HA likely results from a complex inter-relationship between all the factors (Pollard, 2006, 2007).

Regarding referees, technology has been implemented over the years to reduce bias and probability of error: the Goal Line Technology (Psiuk et al., 2014), the vanishing spray (Kolbinger & Link, 2016) and the introduction of additional assistant referees (Albanese et al., 2020).

In 2021, UEFA abolished the away goals tiebreaker, which had been in place since the 1965/66 season with the idea to incentivize road teams to attack, reducing home advantage. Based on the recent criticism and statistics from mid-1970s showing dramatically decrease of HA, UEFA President Aleksander Ceferin abolished the away goals rule (UEFA, 2021).

Besides aiming to promote fairness in the game, IFAB has not always been in favour of adding technology to the game, despite its potential to reduce referee errors and home advantage. In fact, in the 1970 annual general meeting, it was stated that "The Board deprecated the emphasis placed in television recordings and television comment which challenged the authority of the referee. It was requested the televisions to refrain from any slow-motion play back which reflected adversely on any decision of the referee". (IFAB, 1970).

1.4. Video Assistant Referee

In 2018, with the aim of making the referee's job easier and reducing the probability of error, VAR was introduced in the Laws of the Game.

The satisfaction among fans is not guaranteed, however, with several evidence of mixed feelings (Kolbinger & Knopp, 2020; Winand et al.,2021; Hamsund & Scelles, 2021 and Scanlon et al., 2022). Nevertheless, it's believed that the average fan is not aware of the truly impact of VAR (Vale, 2023).

VAR constantly makes correct decisions and represents an improvement versus the Referees first guess. In fact, the accuracy of decisions rose from 92.1% to 98.3% (Spitz et al., 2021). VAR continues for the 7th consecutive season since its introduction and it's important to measure its impact in football dynamics, as presented in the Literature Review.

Regarding the Portuguese context, playing time is currently a hot topic and a subject of increasing interest, caused by recent low rankings on playing time and effective playing time. (CIES Football Observatory, 2019, 2021).

Moreover, another topic that concerns football supporters is the competitiveness of the game. La Liga has decreased it after the adoption of the 3-point scoring system (Soto-Valero & Pic, 2019). The 25 seasons between 1997-2022 saw 31 different clubs winning the title in England, Italy, Spain, Germany and France as compared to 48 different clubs winning the title between 1946-1971 (Ramchandani et al.,2018 and Van der Burg, 2023). Decline in competitiveness is also seen in Champions League (Ramchandani et al.,2023). Evidence suggests that a less attractive product might struggle to command a high market value (Plumley et al., 2018).

While existing research has explored the impact of VAR on game dynamics, as detailed in the Literature Rewiew, there are still unexplored areas regarding its influence on specific game metrics, indicating a gap this study aims to address. Based on this comprehensive introduction, it is understood that goals, home advantage, playing time and competitiveness are crucial to football as the leading sport in the entire world. Therefore, this study aims to focus on 4 different Research Questions (RQ) to study the VAR Impact on Portuguese Professional Football on 1st division between 2013/2014 and 2020/2021:

- RQ 1 – VAR Impact on Total Goals
- RQ 2 – VAR Impact on Home Advantage
- RQ 3 – VAR Impact on Playing Time
- RQ 4 – VAR Impact on Competitiveness

In the upcoming chapter, the current literature will be reviewed. Additionally, Materials and Methods will be evaluated, providing an overview of the models employed in this study. Subsequently, the study will delve into the Results and compare them to the existing research, concluding with a comprehensive assessment of VAR's influence on Portuguese Football.

2. Literature Review

2.1. VAR

VAR introduction in Serie A reduced the number of goals, fouls and yellow cards, and in Bundesliga reduced offsides and yellow cards while increasing extra time (Carlos et al., 2019). This study used a Generalized Linear Model (GLM) and Variance Inflation Factor (VIF) measures were assessed. To compare pre-VAR and post-VAR, it was used paired Z-tests.

In the Chinese Super League, the first season with VAR saw a decrease in offsides, fouls and in HA, but playing time increased (Han et al., 2020). This paper used a T-test and standardized mean differences (SDMs) to compare pre-VAR and post-VAR periods and further estimated GLM.

Games with VAR interventions have more playing time than games with no VAR interventions in Spanish La Liga (Errekagorri et al., 2020). The number of goals also increases with VAR interventions, due to its direct association with Penalty Kicks, which might contrast with Carlos et al. (2019), despite the different scope of analysis. Nevertheless, these effects hardly change the game. This research studied 2018/2019 season and used a One-Way ANOVA for differences in the variables between games with zero, one or two VAR interventions. Posterior, Bonferroni's test was used and Cohen's d effect size was also used.

(Lago-Peñas et al., 2021) analysed 760 games (one season pre-VAR and one season post-VAR) in La Liga and concluded that VAR does not dramatically changes football. There were decreases in offsides and a slight increase in the stoppage time, which is consistent with (Carlos et al., 2019) findings. This study used descriptive statistics and a Mann-Whitney U test for comparisons between periods and estimated size effect using ETA squared statistic. Additionally, the study also assessed a GLM. In Turkey football league, for 2018/2019 season, goals and cards decreased and playing time increased. This paper used a Mann Whitney U test for means comparison. (Gürler & Polat, 2021).

Post-VAR FIFA World Cup saw a significant increase in the number of penalties and total playing time and a decrease in offsides (Kubayi et al., 2022). A Mann-Whitney U test was used to compare differences between VAR periods, distinguish different effect sizes and a GLM was fitted for each match performance variable.

The VAR Introduction in Brazil originated a decrease in the number of fouls, yellow cards and offsides in the 2019 season, compared to 2018 season (Meneguite et al., 2022). This study used a T-test for means comparison after testing for normality of the data.

(Holder et al., 2021) used descriptive statistics on two seasons pre-VAR and two seasons post-VAR in both Bundesliga and La Liga and found that home teams tend to be favoured with respect to penalty kicks, red cards and added time. Home Advantage decreased, but remains existing, which sheds a light on other possible factors accounting for it.

A more sophisticated approach used a difference-in-differences (DiD) and a synthetic control method (SCM) to assess the VAR impact in Referee bias among the leagues that firstly introduced it in 2017/2018 season (Kim et al., 2022). Away goals increased, while home goals remain unchanged.

2.2. Purpose of this study

The existing literature about the VAR has highlighted its impact on soccer dynamics. Yet, there is still room for further improvement in this research area.

In fact, the present studies limit their analysis to a few seasons, mostly using one or two seasons pre-VAR and one or two seasons post-VAR (Carlos et al., 2019; Lago-Peñas et al., 2020; Han et al., 2020; Errekagorri et al., 2020; Gürler & Polat, 2021; Holder et al., 2021 and Meneguite et al., 2022), which might not be tracking pre-VAR trends there might exist already on the studied variables. Therefore, a broader analysis should be performed to account for possible pre-VAR trends.

Despite providing a global and advanced approach to the topic, the DiD approach only accounts for a goal analysis when measuring the home advantage effect. This can leave room for several interpretations once home advantage can be measured using different metrics.

VAR Impact on Total Goals (RQ1) assess a straightforward question and is easily compared with the past studies, being more robust accounting for eight seasons, as seen in the next chapter.

VAR Impact in Home advantage (RQ2) will be studied as a multi factor component. In the previous studies, there were several ways to quantify HA.

Some studies define it as scoring more goals at home than away (Marek & Vávra, 2020), while others use ratios of home goals to total goals (Goumas, 2014) or the ratio of home points to total season points (Inan, 2018; Leite, 2017 and Leite & Pollard, 2018). Traditional views, like Pollard's (1986), consider the percentage of home game wins, whereas recent research like Holder et al. (2021) and Han et al. (2020) looks at penalty kicks, red cards and home goals to gauge home advantage. In a paper addressing how to best quantify for home advantage in handball leagues, the authors induced new metrics to quantify this effect, such that HA can reach values over 100% (Matos et al., 2020).

However, none of these approaches is robust when considering both home goals, home wins and home points. For example, a home team that wins narrowly but loses by larger margins can present a skewed view of HA if only goals are considered. Similarly, the use of home wins and points can lead to misleading interpretations, as teams with the same points can have vastly different win-loss records.

Besides using Home Goal Difference, Home Wins and Home Points, this study will also incorporate market betting odds (normalized as probabilities) for home wins as a measure of expected HA. Betting odds, shaped by a multitude of factors including team performance, injuries, and public sentiment, offer insight into the perceived likelihood of a home win. Thus, odds indicating a more competitive match could suggest a diminished HA.

As such, this paper will present a disruptive analysis of HA (RQ2) in the Portuguese League by a composition of four different metrics: (RQ 2.1) Goal Difference between Home teams and Away Teams; (RQ 2.2) Home Wins; (RQ 2.3) Home Points and (RQ 2.4) Betting Odds.

Additionally, extra time increased in Bundesliga (Carlos et al., 2019) and Playing Time increased significantly in the Chinese League (Han et al., 2020) and Turkey league (Gürler & Polat, 2021) and games with VAR interventions accounted for more playing time (Errekagorri et al., 2020), which make sense due to more adding time. World Cup saw an increase in the number of penalties and playing time and a solid decrease in offsides (Kubayi et al., 2022) and Women's world cup also saw its playing time increased (Zhang et al., 2022).

Playing time is subject of increasing interest in Portuguese football, as previously seen. RQ3 examines the impact of VAR on playing time (RQ 3.1) and on the effective of playing time (RQ 3.2). Simply analysing playing time, as previously studies did, does not account for the fact that a longer game with a lot of additional time will naturally have more playing time, but it is not necessarily a better game to watch. In that sense, effective playing time is crucial to study the impact of VAR on games attractiveness. Moreover, this analysis uses a more comprehensive dataset than existing studies, providing a deeper understanding of VAR's effect on playing time.

The current evidence is consistent in the fact that VAR Introduction significantly increases playing time, which might mitigate the fact that weaker teams are shown to be more propense to waste time during games and therefore decrease the playing time of a game (Greve et al., 2019).

Despite being seen as a friendly factor of the game, research on playing time raises some concerns. Longer games naturally cause players to cover more distance (Altmann et al., 2023). Extra time provokes higher conditioning load in following matches (Rampini et al., 2009 and Winder et al., 2018) and the physical performance of players and technical skills decreases throughout the game (Rampini et al., 2009; Castellano et al., 2011; Liam D 2014; Harper et al., 2014; Peñas et al., 2015 and Goodall et al., 2017), having different impacts based on players position (Rey et al., 2020).

A busy calendar, which is often raised as a problem for several coaches, shows a significant reduction in physical performance throughout the season (Russel et al., 2015), which has impacts on team performance (Hägglund et al., 2013). VAR interventions promotes interruptions and game interruptions are linked with decline in match running performance (Linke et al., 2018).

The decline in physical performance throughout the game and technical skills difference between more successful teams and underdogs, may concern more playing time as an unfair approach and to decrease competitiveness.

RQ4 delves into the VAR impact on the Competitiveness, measured with the variable Winner Margin to quantify to which extent are games more/less balanced in the Portuguese League after VAR introduction. This is a topic seen as a concern by fans and to be decreasing in several European leagues (Van der Burg 2023) and attract negative returns to the market value of a league (Plumley et al., 2018).

In this sense, the bigger the winner margin, the more unbalanced a game is and, the smaller the difference, the more competitive the match probably was. The findings from RQ4 aim to shed new light on VAR's broader effects in football, offering insights that have yet to be fully explored in existing literature.

2.3. Statistical approach

Regarding statistical methods, this paper will firstly address traditional descriptive statistics and non-parametric testes to analyse the impact of VAR on various game metrics. This approach is particularly useful for initial exploratory analysis and when dealing with data without assuming a normal distribution. This framework approach is a common practice amongst researchers (Carlos et al., 2019; Han et al., 2020; Lago-Peñas et al., 2020; Gürler & Polat, 2021; Kubayi et al., 2022 and Meneguite et al., 2022). For categorical variables, as Home Wins, a Chi-Square test is assessed as performed by Erikstad & Johansen (2020) when concluding that successful teams were more likely to receive an incorrect penalty compared with their opponents.

Event Studies are further performed to assess the VAR Impact. This economic approach offers valuable insights into the heterogeneous treatment effects of VAR across different post-treatment time periods, shedding light on immediate impacts versus long-term effects like changes in strategic behaviour or rules' interpretations (Cunningham, 2021).

In an event study using Ordinary Least Squares (OLS) regression, the fundamental idea is to estimate two regressions (pre-VAR and post-VAR) and compare the two. This often involves a "segmented regression" with an interaction term capturing the change in the trend of the outcome variable. The design of an event study revolves around capturing changes over time, comparing before- and after-event periods. Ensuring the treatment is the primary factor causing changes is essential (Huntington-Klein, 2021).

To robust the conclusions, GLM will be tested. This statistical method has been used in several studies on football research (Carlos et al., 2019; Lago-Peñas et al., 2020; Han et al., 2020 and Kubayi et al., 2022). A GLM is a flexible generalization of ordinary linear regression that allows for the dependent variable to have a non-normal distribution, which provides a complementary perspective in case of previously failing the Shapiro-wilk test. It's used to model different types of data that follow distributions like binomial, Poisson, or multinomial. A GLM will show the range of variation of the response variable in each family, and the so-called canonical link function associated with each family (Fox, 2015).

Moreover, a Regression Discontinuity Design (RDD) will also be tested. RDD is a quasi-experimental pre-post test design that is used to identify causal effects when random assignment is not possible. It exploits a cutoff or threshold in an assignment variable to distinguish between those who receive the treatment and those who do not (Lee & Lemieux, 2010). The key assumption is that units just above and just below the cutoff are similar in all respects, except for the treatment. The reason RDD is so appealing to many is because of its ability to convincingly eliminate selection bias. This appeal is partly since its underlying identifying assumptions are viewed by many as easier to accept and evaluate (Huntington-Klein, 2021).

RDD has very little background among football research, with only a few papers have address it: Hon & Parinduri (2014) to measure the impact of three-point rule in games; Moffat (2019), to examine whether participation in the group stage of the UEFA Europa League has a causal impact on the performance of teams in national leagues; Reilly & Witt (2021) to identify causal impact of the Scottish Premiership League split on spectator match attendance; and Bouke Klein Teeselink et al. (2023) that used RDD for binary outcomes while studying if being slightly behind increases the likelihood of winning in basketball.

By comparing the results of Non-Parametric Tests, Event Studies, RDD and the GLM, findings are tested on its consistency across different statistical methods. If all these methods suggest that the introduction of VAR had a significant impact on the metric being studied, this will strengthen conclusions. If the results differ, further investigation might be needed to understand why and to determine which model is more appropriate for the data and research question.

3. Materials and Methods

This chapter will give an overview of the dataset and the framework of different analysis performed on this study.

3.1. Data and variables

The dataset that guides this study is composed of eight consecutive seasons - four pre-VAR seasons and four post-VAR seasons. Seasons from 2017/2018 to 2020/2021 represent post-VAR period and account for 306 games each. Seasons from 2013/2014 to 2016/2017 account for pre-VAR period and have 306 games each, expect for 2013/2014 season when Portuguese League had only 240 games each season. One game from 2018/2019 season was deleted due not having all the information available. Therefore, a total of 2381 games remain to analyse.

Home goals and away goals from 2013/2014 to 2020/2021 were assessed from Football Results, Statistics & Soccer Betting Odds Data (n.d.), which is a credible source and used in other studies (McCarrick et al., 2021). Data on Playing Time and Percentage of Playing Time was provided by Federação Portuguesa de Futebol and data regarding betting odds was retrieved from Odds Portal (n.d.), which is also a credible data source.

The final dataset has 14 metrics on a game-level granularity. The chapter 4.1 Descriptive Statistics shows detailed descriptive statistics for these variables, as a starting point for further analysis.

VARIABLE	MEANING	TYPE
Year	Year of each season (eg: 2014/2015 is Year 2015)	Nominal
Treatment	Either a game with (1) or without (0) VAR	Binary
Home Goals	Goals scored by the home team	Continuous
Away Goals	Goals scored by the away team	Continuous
Total Goals	Total goals scored in the game	Continuous
Home Advantage	Goal difference between home and away	Continuous
Home Win	Either a home win (1) or not a home win (0)	Binay
Home Points	Points for the home team based on result (0,1 or 3)	Ordinal
Winner Margin	How much more goals did the winner score	Continuous
Odd Home	Normalized probability for home team to win	Continuous
Odd Draw	Normalized probability of a draw	Continuous
Odd Away	Normalized probability for away team to win	Continuous
Playing Time	Total Playing Time of the game	Continuous
Effective playing time	Playing Time as percentage of total game time	Continuous

Table 3.1 Dataset Variables

The variable Year states the season of the game, being that 2018 corresponds to 2017/2018 season. Treatment refers to the game having VAR or not, i.e., if the game is prior or after the season 2017/2018. Home and Away goals refer each team's score and Total Goals is the sum of both. Home Advantage, as traditional assessed by researchers, measures the goal difference between home and away teams. Home Wins accounts either for 1 or 0, whether the home team wins the game and Home Points refers to the total points awarded by that result. Odd home, odd draw and odd away refer to the normalized probabilities for the respective outcome. Playing time refers the total time that the ball was in play and effective playing time measures the proportion of this playing time according to the game duration.

3.2. Statistical analysis

3.2.1. Descriptive Statistics

To start the statistical analysis, we'll first look closely at each variable and assess its descriptive statistics, accounting for mean, variance, median, quantiles and several visualizations. Furthermore, by using a correlation matrix, we'll check if the variables are too related to each other, possibly causing multicollinearity issues. After the variable selection, VIF is measured to help strengthen the previous conclusions.

3.2.2. Non-parametric tests

The variables in the dataset were tested for normality using a Shapiro Wilk Test and all were considered to not following a normal distribution ($p < 0.05$). Therefore, after the initial approach, the study addresses non-parametric tests, due to its easiness of use once they do not require any specific data distribution when compared with traditional parametric tests, as T-test.

Despite most of the teams being the same in the two VAR periods (pre and post), players and coaches and even teams change year to year, as do game dynamics and tactics, so that pre-VAR and post-VAR are considered to be two independent groups. Therefore, a Mann Whitney U Test will be performed to compare the two groups for different game variables. For metrics as Home Wins (RQ 2.2), which is binary, a Chi-Square test is performed and compared with a Fisher's Exact test for robustness purposes.

These results are posteriorly evaluated by a Cohen's d, which incorporates standard deviation to measure effect size and assess the magnitude of the changes.

3.2.3. Event Studies

Event Studies (ES) are traditionally used in economics to assess the impact of a specific event on the value of a firm, for instance. The basic idea of an Event Study is that, at a certain time, an event occurred, leading a treatment to be put into place at that time. Whatever changed from before the event to after is the effect of that treatment (Huntington-Klein, 2021).

ES will be applied in this study assuming that VAR introduction is a treatment put into place in 2017/2018 season in Portugal 1st division. As such, changes in game dynamics pre-VAR to post-VAR represent the effect of VAR introduction.

ES commonly uses OLS regressions to estimate returns and therefore one ES will be conducted for each individual research question, considering every control variable: Total Goals, Home Advantage, Home Wins, Home Points, Team strength, Playing Time, Effective Playing Time and Winner Margin. After performing the individual ES, a comprehensive model with all the control variables is performed. Note that, as further seen in chapter 4.1 Descriptive Statistics, not every control variable will be added simultaneously to the model, so that regressions are free from possible multicollinearity issues. Moreover, for binary variables as Home Wins, a logistic regression will be performed instead of OLS regression.

The discussion of these regressions is found in the following chapter and their output table can be seen in the Appendix B - Event Studies.

$$Y_i = \beta_0 + \beta_1 Treatment_i + \varepsilon_i \tag{3.1}$$

$$Y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Variable1_i + \varepsilon_i \tag{3.2}$$

$$Y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Variable2_i + \varepsilon_i \tag{3.3}$$

...

$$Y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Variable1_i + \ldots + \beta_n Variable\ n_i + \varepsilon_i \tag{3.4}$$

The first event study approach (3.1) will simple regress the dependent variable assessing the direct impact of VAR. Moreover, each control variable will alternatively be added to the model, until a comprehensive model is regressed with all the control variables (3.4).

From the equations, the dependent variable Y_i is the variable being predicted depending on the research question. β_0 is the intercept (constant) term and β_1 to β_n are the coefficients of the predictor variables that represent the relationship between each predictor and the dependent variable. $Variable_1$ to $Variable_n$ are the predictor variables (Total Goals, Home Advantage variable, Home Odd, Playing Time, and Winner Margin) used in the model and ϵ represents the error term, which accounts for the variability in the dependent variable that is not explained by the predictors and it is assumed to be normally distributed with mean 0 and constant variance.

Furthermore, the procedure is replicated incorporating year fixed effects, which is a modelling approach used to account for unobserved or time-invariant factors that may be influencing the outcome variable.

In the year fixed effect models, the treatment variable is not directly included in the models. However, its effects are still observable, once it is strictly predicted by year (pre-2018 = 0, post-2018 = 1). This way, fixed effects allow to examine the year-to-year fixed effects without possible issues derived from Treatment being perfectly predicted by years.

$$Y_{it} = \beta_0 + \beta_1 Variable1_{it} + \sum_{t=1}^{n} yi\ \mathrm{Year}_i + \varepsilon_{it} \tag{3.5}$$

$$Y_{it} = \beta_0 + \beta_1 Variable1_{it} + \beta_2 Variable2_{it} + \sum_{t=1}^{n} yi\ \mathrm{Year}_i + \varepsilon_{it} \tag{3.6}$$

$$Y_{it} = \beta_0 + \beta_1 Variable1_{it} + \beta_2 Variable3_{it} + \sum_{t=1}^{n} yi\ \mathrm{Year}_i + \varepsilon_{it} \tag{3.7}$$

...

$$Y_{it} = \beta_0 + \beta_1 Variable1_{it} + \beta_2 Variable2_{it} + \ldots + \beta_n Variable\ n_{it} + \sum_{t=1}^{n} yi\ \mathrm{Year}_i + \varepsilon_{it} \tag{3.8}$$

Similar to the previous approach, year fixed effects are performed on a growing complexity, starting with treatment alone (3.5) and reaching a final model which accounts for all the control variables (3.8). In these regressions, β_0 is the intercept and β_1 is the coefficient for the main independent variable $Variable\ n_{it}$, which is the main independent variable for observation i in year t. N is the number of fixed year effects (dummy variables for each year) and γt are the coefficients for the fixed year effects. The summatory means the sum of year fixed effects for each year from 1 to n and ε_{it} is the error term for observation i in year t, representing the unexplained variation in the dependent variable.

To conclude Event Studies approach, interaction terms are added to the model 6, which is a simple but robust model, avoiding concerns of over-complex models when adding interaction terms to a fixed effects model. The interaction terms allow to examine how the relationship between two or more variables changes based on the presence of another variable. In the study of the impact of VAR in football, interaction terms can be valuable for exploring how the effects of VAR may vary under different conditions or in the presence of other variables.

$$\begin{aligned} Y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Variable1_i \\ + \beta_3 Treatment_i \times Variable1_i + \ldots + \beta_{n+1} Variable\ n_i \\ + \beta_{n+2} Treatment_i \times Variable\ n_i + e_i \end{aligned} \tag{3.9}$$

In this model (3.9), the addition of $\beta_n Treatment_i \times Variable\ n_i$ represents the interaction between treatment and variables, which is useful in cases where the effect of one predictor on the outcome is influenced by VAR.

Event Studies, however, raise concerns regarding autocorrelation. Therefore, these regressions use heteroskedasticity and autocorrelation (HAC) robust standard errors. To further check on this, after each regression a Durbin-Watson test is performed to detect the presence of autocorrelation in the residuals.

3.2.4. Robustness Checks

To robust and complement this study framework, a Generalized Linear Model will be performed. GLM is a flexible generalization of linear regression models once it allows for nonlinearity. Simply put, the GLM take the regression model and pass it through a function to make the prediction. This function is called a link function and allows GLM to work with several distributions when OLS regression do not hold, once is adaptable to the dependent variable properties (Huntington-Klein, 2021). GLM are commonly used in this research field, as performed by Carlos et al. (2019), Lago-Peñas et al. (2020), Han et al. (2020) and Kubayi et al. (2022).

GLM are a plausible complement based on the concerns on the normality of data since GLM assumes that the relationship between dependent and independent variables may be non-linear. There are three components to a GLM: a Random Component, which refers to the probability distribution of the response variable (Y); a Systematic Component, referring to the explanatory variables $(X_1, X_2, \ldots X_k)$ as a combination of linear predictors; and a Link Function ($g(\mu)$) which specifies the link between random and systematic components, i.e., how the expected value of the response relates to the linear predictor of explanatory variables.

$$Y_i = \beta_0 + \beta X_i + \epsilon_i \qquad \text{or} \qquad E(Y_i) = \beta_0 + \beta X_i \tag{3.10}$$

In Simple Linear Regressions (3.10), the random component (Y) is a response variable and has a normal distribution, and generally is assumed $\epsilon_i \sim N(0, \sigma^2)$. The systematic component X is the explanatory variable (continuous, discrete or both) and are linear in the parameters $\beta_0 + \beta X_1$. The link function is called the identity link: $\eta = g(E(Y_i)) = E(Y_i)$.

In Binary Logistic Regression, the random component (Y) assumes a binomial distribution, the systematic component X are explanatory variables and linear in the parameters $\beta_0 + \beta X_1$. The link function assumes a logit expression: $\eta = logit\ (\pi) = log\ (\frac{\pi}{(1-\pi)})$.

Regarding Loglinear Models, they model the expected cell counts as a function of levels of categorical variables. As such, the random component (Y) assumes a Poisson distribution, and the link function is expressed as $log\ (\mu_i) = \beta_0 + \beta_1 X_1$ or $\mu = e^{(\beta_0 + \beta_1 X)}$, where μ is the predicted value of Y given X; exp(β_0) is the effect on the mean of μ when X = 0 and exp(β_1) is the multiplicative effect on the mean of Y for a one-unit increase in X. e represents a constant value of approximately 2.72.

A Poisson distribution is used to characterize count data, as the variable Goals. One of the major assumptions from Poisson distributions is that the mean and variance are equal. If this assumption is not met, a Negative Binomial (NB) model should be used, which is similar to Poisson model, but incorporates an additional term to account for the excess variance. NB is then used to model Home Points variable. A Zero Inflated Poisson model is used when the number of zeros is considerably large, which is the case of Winner Margin. Finally, a Linear Regression is used to model Home Advantage (Home Goal Difference), Home Odds, Draw Odds and Away Odds, Playing Time and Effective Playing Time. Home Advantage was not model using a Poisson distribution once it accounts for negative values, which is not a valid assumption for a Poisson distribution.

Overall, bringing GLM to the analysis provides a more complete approach due to its wide range of applications, which improves the analysis, once Event studies are mostly used in finance and economics. GLM results will be presented in Results section and their regression outputs can be found in the Appendix C - Generalized Linear Models.

Additionally, a Regression Discontinuity Design will also be used to robust this study. A RDD is a quasi-experimental design used to estimate causal effects when a treatment is assigned based on a continuous score. In the context of this study, RDD can help assessing the immediate impact of VAR by comparing outcomes just before and after the introduction threshold, assuming no other confounding factors change between (Huntington-Klein, 2021). RDD compares the game metrics on either side of the cutoff treatment, assuming they would be similar if there was no VAR introduction so that, if they are different, one can probably attribute the difference to whatever has happened at that cutoff – VAR Introduction. RDD has very little background among football research, as seen in Literature Review, with only a few authors addressing it: Hon & Parinduri (2014), Moffat (2019), Reilly & Witt (2021) and Bouke Klein Teeselink et al. (2023). As such, this study offers a new perspective to study different game metrics and assessing VAR impact using RDD.

A visual representation of the cutoff for each research question is presented in the Results section, so that it either robust or contrast with the previous analysis. Besides discontinuities, RDD also uses OLS regression analysis, which could be exemplified as follows:

$$Y = \beta_0 + \beta_1(Running\ -\ Cutoff)\ +\ \beta_2 Treated + \beta_3(Running\ \ -\ Cutoff) * Treated\ + \varepsilon \tag{3.11}$$

The equation (3.11) is a simple linear approach to regression discontinuity, where Running is the running variable (variable Year), which is centered around the cutoff by using (Running – Cutoff), which takes a negative value for pre-VAR, zero at the cutoff and positive value post-VAR. Treated is an indicator for being treated (either prior or post-VAR), which also means above of below cutoff. RDD can be applied on all the data or limited to a bandwidth around the cutoff, which enhances the focus on a shorter-term analysis (smaller bandwidths) or wider analysis (bigger bandwidths).

Most of the RDD literature, however, focusses on continuous outcomes, with a scarce focus on binary outcomes. In this study, one of the ways to measure Home Advantage is through Home Wins with a binary outcome. Although many empirical studies using RDD consist of an outcome variable that is continuous where OLS is the modelling strategy of choice, the RDD can easily be incorporated with just about any type of regression analysis such as logistic regression (Lesik, 2008), as such:

$$\log(HomeWin) = \beta_0 + \beta_1 Treatment\ +\ \beta_2(Running\ -\ Cutoff) \tag{3.12}$$

$$\log(HomeWin) = \beta_0 + \beta_1 Treatment\ +\ \beta_2(Running - Cutoff) + \beta_3 Treatment * (Running\ -\ Cutoff) +\ \beta_4(Running\ -\ Cutoff)^2 + \beta_5 Treatment * (Running\ -\ Cutoff)^2 \tag{3.13}$$

Equation (3.12) specifies the relationship between the assignment predictor and the outcome variable as being linear once the assignment predictor is specified to the first power. According to Lesik (2008), a possible way to assess the functional form specification of RD model is to include higher order non-linear terms along with their respective interactions to the model to see if the addition of these terms has a significant impact on the estimate of the treatment effect. Therefore, both regressions (3.12) and (3.13) will be performed to access the VAR Impact on Home Wins. Their visualizations can also be found in the Results section, and their regression outputs can be seen in Appendix D - Regression Discontinuity Design.

All in one, GLM and RDD offer a complement to this study on multivariate analysis and a grasp on causal inference, respectively, when assessing the VAR impact on Portuguese Football. The results will be further presented in the Robustness Checks section.

4. Results

This chapter provides results for the framework assessed in this study. Chapter 4.1 provides descriptive statistics based on the variables considered for the research questions and Chapter 4.2 performed a non-parametric test to measure the differences between pre-VAR and post-VAR games. Moreover, Event Studies results are presented, with detailed analysis of its approaches based on simple models, fixed year effects and interaction terms. Additionally, these results will be complemented with GLM and RDD methods in the Chapter 5.

4.1. Descriptive Statistics

The figure A.1. shows a raw correlation matrix, in which the variables Home Goals and Away Goals demonstrate significant correlations (>70%) with Home Advantage (as Home Goal Difference). HA also shows notable correlations of 78% with Home Wins and 85% with Home Points. In the same direction, Home Points and Home Wins are 96% correlated. Playing Time is strongly correlated with Effective Playing Time at a level of 82% and the normalized betting odds for both home and away wins display a remarkably high correlation of 95%.

Each research question will have its own comprehensive regression model, so that control variables will be ad hoc chosen based on statistical significance and relevance to the models. As such, Home Advantage will be represented in the models by one of the high related metrics (Home Advantage, Home Wins or Home Points), either Playing Time or Effective Playing Time will be used and Teams Strength will be measured by either Home Odd or Away Odd, according to which variable at the time adds more explanatory power without compromising the model.

Therefore, the below figure shows a possible final correlation matrix, based on the up mentioned assumptions. To address multicollinearity concerns, VIF was measured for this reduced dataset. The variables have VIF values close to 2, which indicate lower multicollinearity.

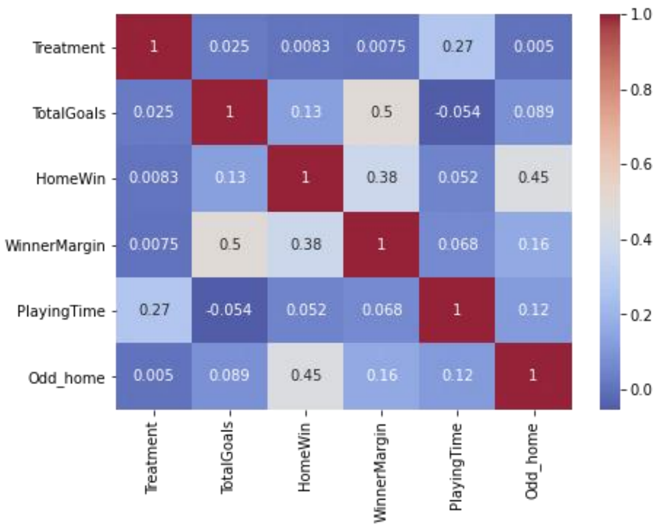

Figure 4.1. Final Correlation Matrix

The below descriptive statistics are complemented in the Appendix A - Descriptive Analysis showing data descriptive statistics, as the mean, variance, min, max and quantile information for all the variables both prior and after the VAR Intervention.

4.1.1. RQ 1 - VAR Impact on Goals

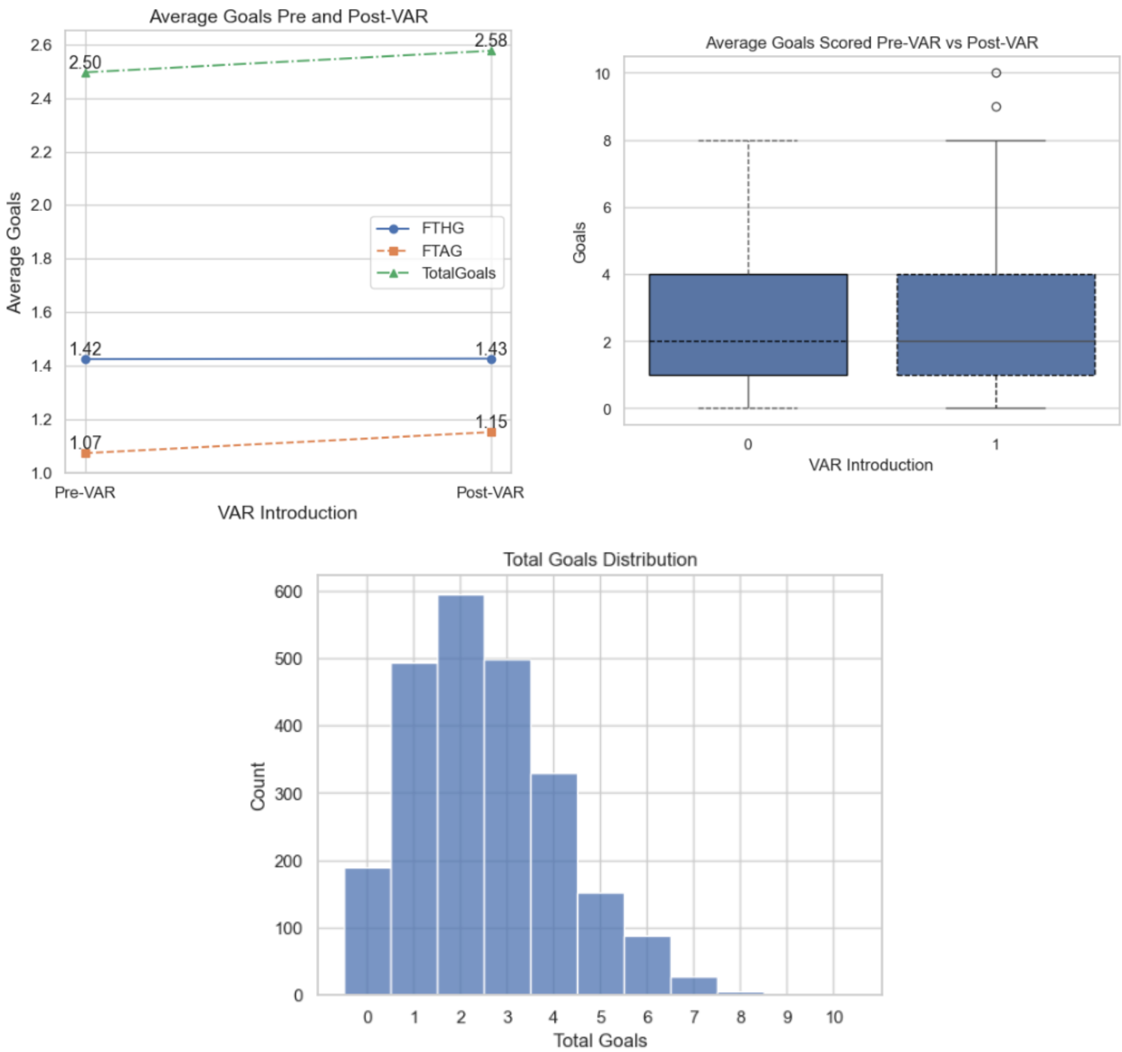

Figure 4.2. RQ 1 - Descriptive Statistics

The away team used to score, on average, 1.07 goals per game pre-VAR and now they score 1.15 goals per game. The post-VAR period saw an increase of 7% in the Away Goals. Home goals remained roughly the same, which results in an increase of 3% in Total Goals. There was an increase of 0,08 goals per game, that represents one increment goal in each 12,5 games. While this value might seem low, it can make a difference on close games for the away teams, considering the vast majority of this increase is due to away goals.

Regarding the boxplots, there are no major differences between pre and post-VAR games, only an increase in outliers in post-VAR games. Moreover, considering that the mean and the variance are only slightly different (see Appendix A - Descriptive Analysis), the histogram figure suggests that Total Goals follows a Poisson distribution, which will be taken into consideration on the next chapters.

4.1.2. RQ 2- VAR Impact on Home Advantage

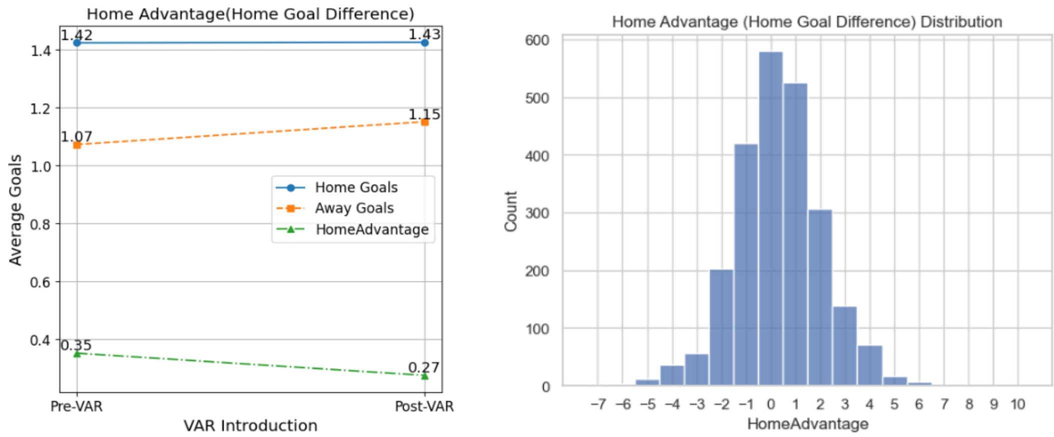

Figure 4.3. RQ 2.1 - Descriptive Statistics

Assessing HA as the goal difference between home teams and away teams, the HA decreased 0.08 goals per game from pre-VAR to post-VAR. Once again, this decrease is mostly due to the increase in goals scored by the Away Team.

Besides the suggestion of figure 4.3, the variable HA (as Home Goal Difference) does not follow a gaussian distribution, as mentioned in the next chapter. However, this variable differs from Total Goals, once it accounts for negative values, which will be taken into account when using robustness methods as GLM.

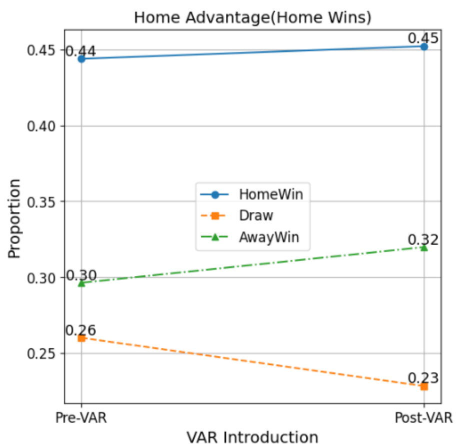

Figure 4.4. RQ 2.2 - Descriptive Statistics

When considering the proportion of Home Wins, this value rose from 44% to 45%. Away Teams win, on average, more 8% of the games than they did pre-VAR. Tied games are less common after VAR Introduction, decreasing around 3 percentage points.

Home Wins is a binary variable, which needs to be accounted for during the applied methods on the following chapters.

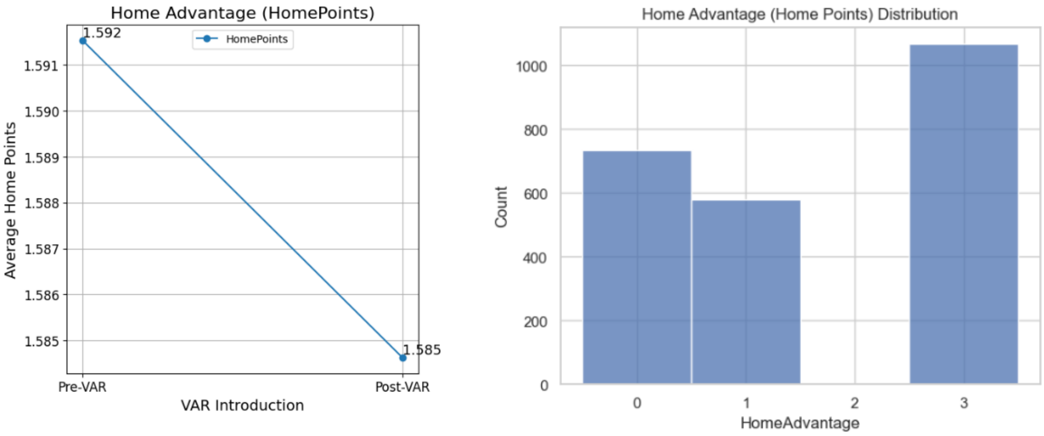

Figure 4.5. RQ 2.3 - Descriptive Statistics

Despite this idea that home wins slightly increase, Home Teams are seeing less points staying at home. In fact, draws are considerably less common since VAR Introduction (figure 4.4) and Home Teams averaged a very tiny decrease in points gained at home (0.4%).

The variable Home Points appears to be over dispersed, as suggested by its difference between the mean (pre-VAR: 1.59 and post-VAR: 1.58) and the variance (pre-VAR: 1.72 and post-VAR: 1.79).

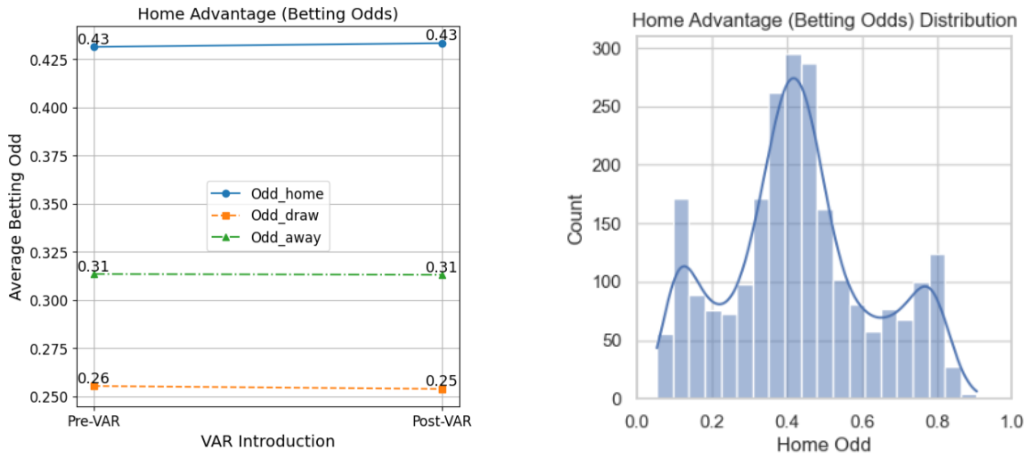

Figure 4.6 RQ 2.4 - Descriptive Statistics

Regarding betting odds, which were treated and normalized as probabilities, HA seem to be relatively steady, with no major changes in odds.

The average probability for a home win both pre-VAR and post-VAR is 43% which indicates no major change with VAR Introduction. It is important to note that probabilities are accounted for having 3 possible results (home, away and draw) and, therefore, a 43% chance of win intrinsically indicates a home advantage, once the average away probability to win a game is around 31%.

4.1.3. RQ 3- VAR Impact on Playing Time

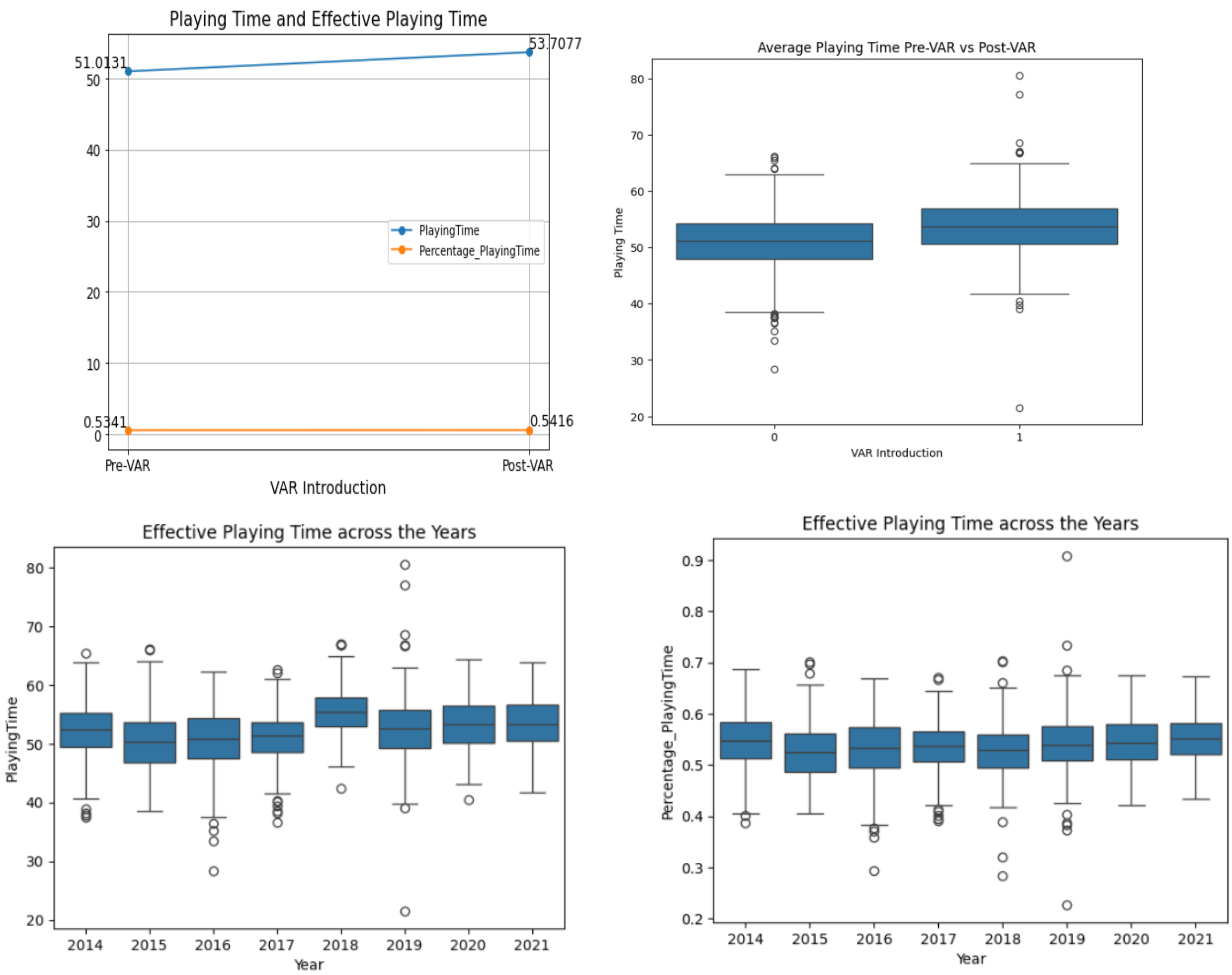

Figure 4.7. RQ 3 - Descriptive Statistics

Post-VAR Playing Time increased in 5%, which equals around 2,40 minutes. However, despite seem a big increase, it only resulted in an increase of 0.75% on the effective playing time. This means that games are getting bigger playing times and the percentage of playing time is also growing, although at a lower level, which might induce future research on marginal gains of additional time.

Despite the quantile amplitude of playing time have slightly decreased, post-VAR games show a bigger discrepancy in outliers, having smaller outliers (the minimum playing time was around 22 minutes) and also bigger outliers (the maximum playing time was around 80 minutes), which might due to longer VAR interventions.

Although the playing time had an immediately increase after VAR introduction in 2018, the effective playing time actually decreased, only reaching 2014 levels in 2020. This suggests that longer playing time is not necessarily more appealing to the fans.

4.1.4. RQ 4- VAR Impact on Competitiveness

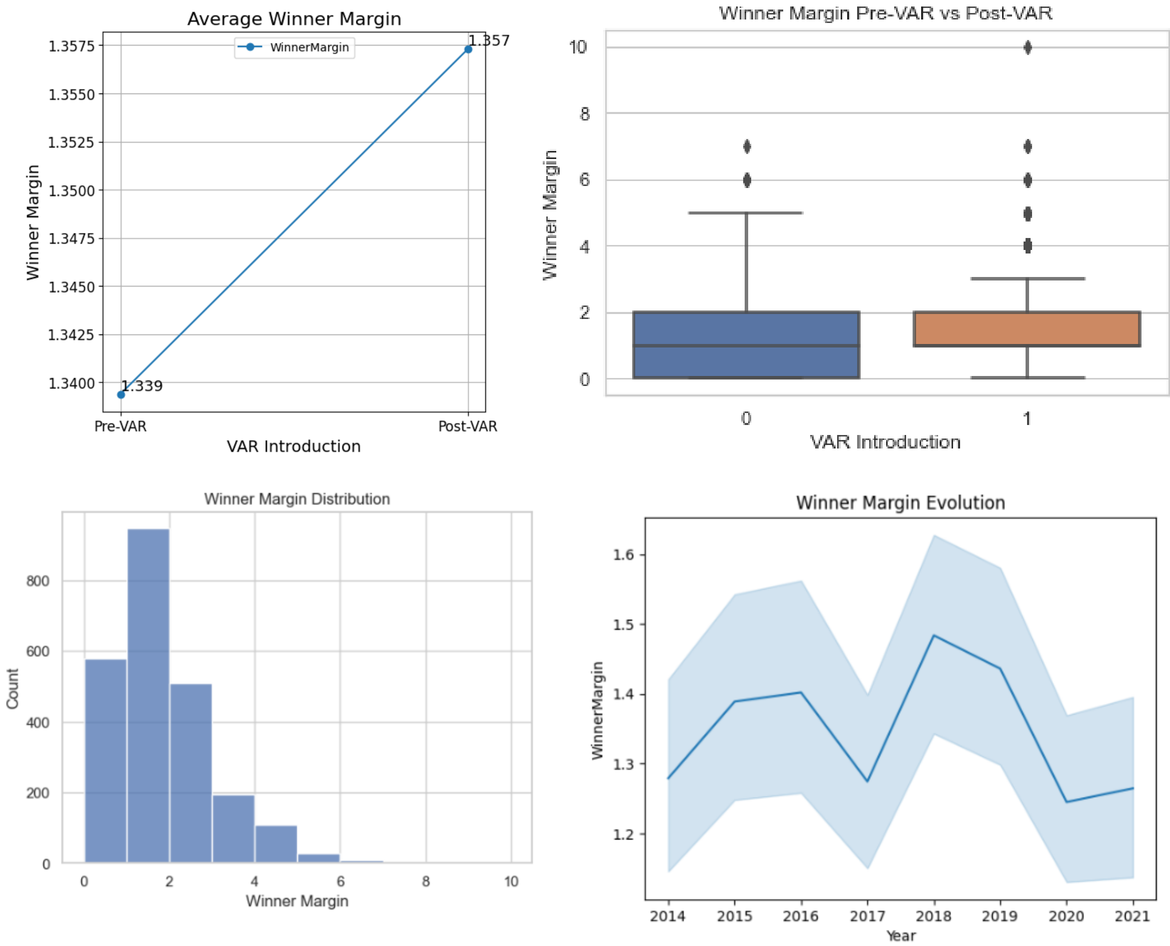

Figure 4.8. RQ 4 - Descriptive Statistics

Winner Margin, which measures to which extent is a game score close, rose around 1.34% in post-VAR era, which means that games got slightly less balanced. However, this is a very tiny difference that requires further confirmation. The line graph suggests the initial years of VAR may have increased the Winner Margin, with a post decrease in Covid-19 era.

Pre-VAR boxplot indicates a wider spread of data, which suggests greater variability in winner margins before VAR was introduced.

Finally, the histogram indicates that games are quite often tied, which is the second most common value with 24%. This fact will be further remembered when choosing the link function for GLM in chapter 5.

4.2. Mann-Whitney U Test

The comparison between no VAR and VAR seasons were run using the Mann-Whitney U test due to the non-normal distribution of match statistics (Shapiro-Wilk test for all variables produced values with $p < 0.05$). For Home Wins, a binary variable, both Chi-Square and a Fisher's exact test were performed and results presented.

Cohen's d was also measured, being that values close to 0.20 is indicative of a modest effect size, a Cohen's d near 0.50 signifies a moderate effect size, denoting a clearer distinction between the groups that may carry practical significance and,lastly, a Cohen's d of 0.80 or above reflects a robust effect size, pointing to a pronounced and practically significant divergence between the group averages.

VARIABLE	MEAN PRE-VAR	MEAN POST-VAR	T STAT	P VALUE	COHEN'S D	CI LOW COHEN'S D	CI HIGH COHEN'S D
Total goals	2.4965	2.5773	696042.5	0.4630	0.0502	-0.0302	0.1305
Home goals	1.4240	1.4260	707185	0.9541	0.0016	-0.0788	0.0820
Away goals	1.0725	1.1513	685621.5	0.1591	0.0687	-0.0116	0.1491
Home advantage	0.3515	0.2747	721576	0.4142	-0.0431	-0.1235	0.0372
Home points	1.5915	1.5846	713238	0.7428	-0.0052	-0.0856	0.0752
Home win (Fisher's)	0.4439	0.4522	0.1338[1]	0.7146	0.4070	-0.064	0.0953
Home Win (Chi-Square)	0.4439	0.4522	1.0258[2]	0.6604	0.0167	-0.0637	0.0970
Odd home	0.4313	0.4332	705552.5	0.8785	0.0099	-0.0704	0.0903
Odd draw	0.2553	0.2538	0.6218	0.5342	-0.0255	-0.1059	0.0549
Odd away	0.3134	0.3130	708666.5	0.9739	-0.0023	-0.0827	0.0781
Playing time	51.0131	53.7077	483588	0.0000	0.5703	0.4899	0.6506
Effective playing time	0.5341	0.5416	653718	0.0012	0.1427	0.0624	0.2231
Winner Margin	1.3394	1.3573	700896	0.6521	0.0149	-0.0654	0.0953

Table 4.1 Mann Whitney U test, Chi-Square and Fisher's Exact Test

[1] Fisher's Exact Test accounts for odds ration instead of Mann Whitney U stat
[2] Chi-Square accounts for a Chi square metric instead od Mann Whitney U stat

4.2.1. RQ 1 – VAR Impact on Goals

The mean slightly increased from pre-VAR to post-VAR, but the Mann Whitney U test shows this difference is not statistically significant (p=0.4630). The effect size is small (Cohen's d=0.0502) and indicates a most likely negligible impact of VAR on total goals scored.

In the context of the research, this analysis would suggest that while there is a small increase in the number of Total Goals scored in the post-VAR period, it cannot confidently be attributed to the introduction of VAR. Even though, a difference of 0.08 goals per game may still be meaningful, especially in the context of football where small changes can have an impact.

4.2.2. RQ 2 – VAR Impact on Home Advantage

As already seen, the statistical analysis shows that the number of goals scored by home teams remained largely unchanged after the introduction of VAR. In contrast, away goals increased by 0.08 goals per game, which would mean around 1 goal for each 13 games. Therefore, the HA (as goal difference) decreased, despite not being significant. The comparison of home points earned before and after the introduction of VAR also decrease but not at a statistically level.

Overall, the away goals increased, home goal difference decreased and home points decreased, which states a decreased in the overall home advantage, however being non-significant. Besides, away goals, home wins (both for Chi-Square and Fisher's exact test) and probability for a home win had a slightly non-significant increased, which induces the conclusion that VAR had no major impact on Home Advantage.

4.2.3. RQ 3 – VAR Impact on Playing Time

The introduction of VAR has led to a statistically significant extension in playing time, with a p-value of less than 0.0001 indicating a highly reliable result. The magnitude of this increase is considerable, as reflected by a moderate to large effect size, with Cohen's d calculated at 0.5703.

Moreover, effective playing time has also seen a significant uptick, as denoted by a p-value of 0.0005, and the effect size of 0.1427 points, small to moderate in scope.

These findings suggest that the implementation of VAR is associated with a noticeable augmentation in both the overall playing time and effective playing time, which are very good news to Portuguese football, despite some concerns that will be further discussed.

4.2.4. RQ 4 – VAR Impact on Competitiveness

The high p-value indicate that possible changes on Winner Margin are not statistically significant. In terms of game balance, this suggests that the introduction of VAR has not had a substantial impact on how balanced games are in terms of scoring.

4.3. Event Studies

The following event studies used HAC standard errors as an adjustment to account for possible heteroskedasticity and autocorrelation in the error terms. The output regression tables can be seen on Appendix B - Event Studies.

VIF scores previously assessed fell well below the threshold so that this regression model does not have significant multicollinearity issues.

4.3.1. RQ 1- VAR Impact on Goals

Model 1, as presented in Materials and Methods, is the simplest model and serves as a baseline to assess the direct impact of VAR. While there is an observed increase in the goals with the introduction of VAR, it is not statistically significant.

$$TotalGoals_i = \beta_0 + \beta_1 Treatment_i + +\beta_2 TeamStrenght_i + \beta_3 HomeWin_i + \beta_4 WinnerMargin_i + \beta_5 PlayingTime_i + \varepsilon_i \quad (4.1)$$

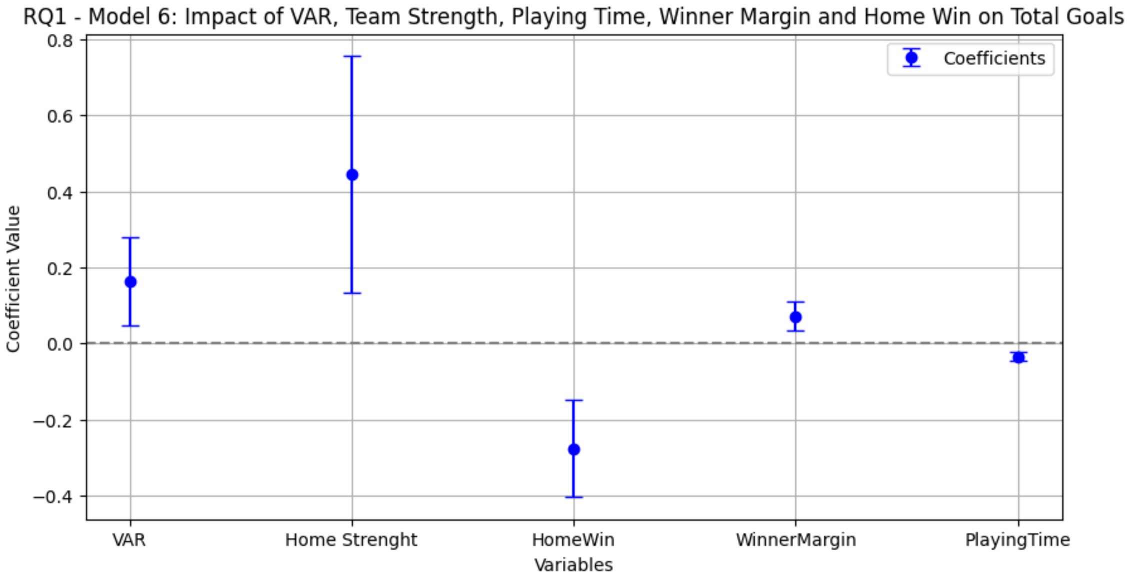

Figure 4.9. RQ 1 - Event Study (Model 6)

The model 6 (4.1) is a comprehensive model and shows a significant positive coefficient for Treatment, indicating that, when controlling for all the variables, VAR introduction is associated with an increase in the total number of goals.

This model was tested for autocorrelation with a score of 2.04, which suggests that there is little to no autocorrelation in the residuals. Moreover, this model has the highest explanatory power among all models (26.4%) and suggests that VAR may play a role in the total goals scored when considering multiple factors together.

VAR introduction is statistically significant ($p < 0.01$), suggesting a robust impact on Total Goals when several control variables are considered. Home Team Strength and Winner Margin (alongside VAR) contribute positively for more goals in a football match. In contrast, Playing Time and Home Win account for fewer goals in a game. The negative impact of Playing Time is consistent with the other outputs and might suggest that players might be tired and aren't so effective, as previously seen. Moreover, it might be connected with the time wasted on goals celebration, which decreases playing time (Fisher, 2023).

To complement the analysis, this regression was controlled for yearly variations and capture time trends. The same models were performed again, but this time with fixed year effects.

$$TotalGoals_{it} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 HomeWin_{it} + \beta_3 WinnerMargin_{it} + \beta_4 PlayingTime_{it} + \sum_{t=2014}^{2021} yi \ \ Year\ t \ + \varepsilon \tag{4.2}$$

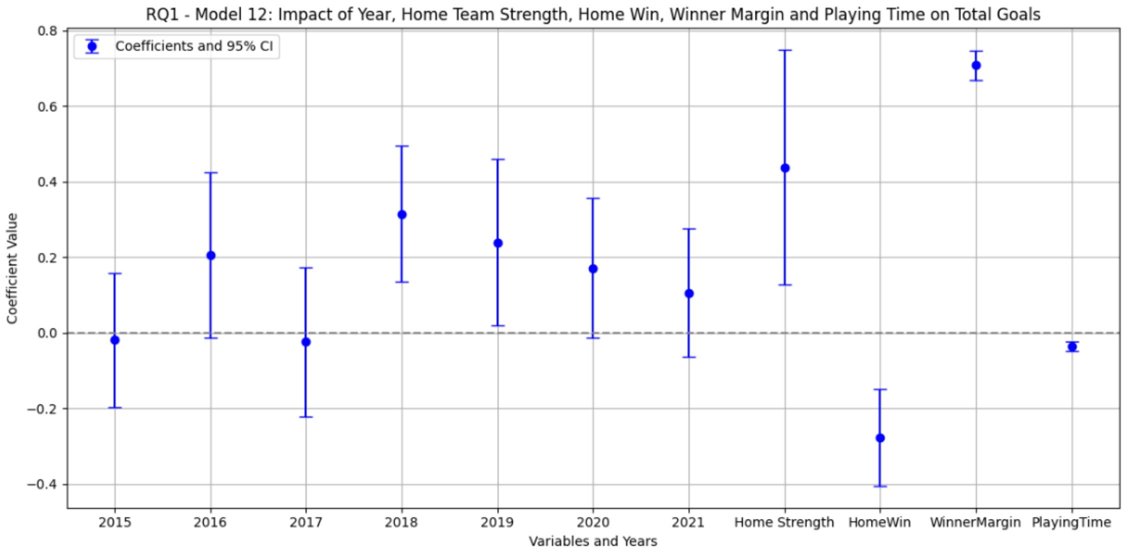

Figure 4.10. RQ 1 - Event Study (Model 12)

The significant coefficient for 2018 ($p < 0.01$), 2019 ($p < 0.05$) and 2020 ($p < 0.1$) in model 12 (4.2) could be interpreted as capturing the initial impact of VAR introduction. This suggests that the introduction of VAR is associated with an observable increase in total goals in its inaugural years. However, the significant coefficient ($p < 0.1$) of 2016 might indicate that the increase in goals might have started earlier.

Subsequent years in the model 12 (4.2) do not show a consistently significant effect as strong as the earlier years, indicating that the initial impact of VAR may have stabilized or that other factors in those years may have influenced goal scoring, which is consistent with literature on Covid-19.

It is also worth note that coefficients for Team Strength, Playing Time and Winner Margin remained consistent throughout the models. Home Win, however, changed from significantly positive to significantly negative, which warrants caution on conclusions for Home Wins.

$$\begin{aligned} TotalGoals_i = \beta_0 &+ \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\ &+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 PlayingTime_i \\ &+ \beta_5 Treatment_i \times PlayingTime_i + \beta_6 HomeWin_i \\ &+ \beta_7 Treatment_i \times HomeWin_i + \beta_8 WinnerMargin_i \\ &+ \beta_9 Treatment_i \times WinnerMargin_i + e_i \end{aligned} \tag{4.3}$$

The introduction of interaction terms did not substantially improve the model fit. In fact, the VAR Introduction becomes non-significant. This discrepancy might be due to the interaction terms absorbing some of the effects of VAR.

Home Win and Playing Time have a consistently negative and significant impact on total goals in models 6 (4.1) ,12 (4.2) and 13 (4.3), suggesting that games won by the home team are associated with fewer total goals and that longer playing times are associated with fewer goals scored.

The coefficient for Home Strength and Winner Margin is positive and significant in the three models, indicating that stronger home teams are associated with an increase in total goals and that larger winning margins are associated with more total goals, which is expected.

The R^2 values are nearly identical, and the models have a significant F-statistic, indicating they both explain a similar amount of variance in total goals and are a good fit for the data.

The significance of the VAR introduction in models 6 and 12 suggests that VAR may have an impact on total goals. However, this effect seems to be sensitive to model specification since it is not significant when interaction terms are included.

4.3.2. RQ 2- VAR Impact on Home Advantage

4.3.2.1. Home Advantage as Home Goal Difference

In the simplest model – model 1, the coefficient for the VAR Introduction is negative, suggesting that the introduction of VAR is associated with a slight non-significant decrease in HA.

$$\begin{aligned} HomeAdvantage_i &= \beta_0 + \beta_1 Treatment_i + +\beta_2 TeamStrenght_i + \beta_3 TotalGoals_i \\ &+ \beta_4 PlayingTime_i + \beta_5 WinnerMargin_i + \varepsilon_i \end{aligned} \tag{4.4}$$

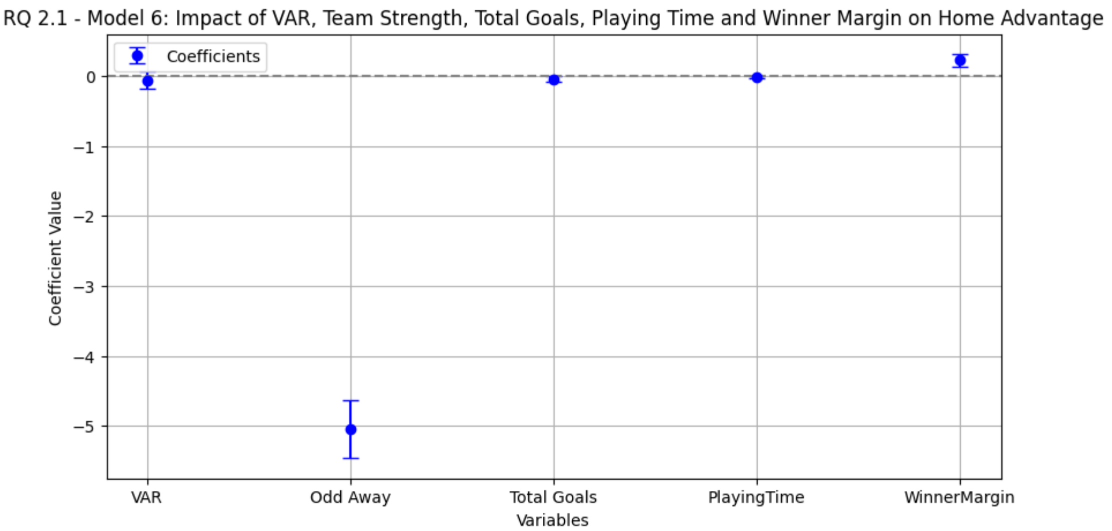

Figure 4.11. RQ 2.1 – Event Study (Model 6)

Model 6 (4.4) was tested for autocorrelation with a score of 1.96, which suggest that there is little to no autocorrelation in the residuals. This model explains about 31% of the variability in HA and the coefficient for Treatment is -0.0810, although it is not statistically significant. This consistent finding across various models reinforces the conclusion that the introduction of VAR does not have a statistically significant impact on HA.

$$HomeAdvantage_{it} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 TotalGoals_{it} + \beta_3 WinnerMargin_{it} + \beta_4 PlayingTime_{it} + \sum_{t=2014}^{2021} yi\ Year\ t\ + \varepsilon \tag{4.5}$$

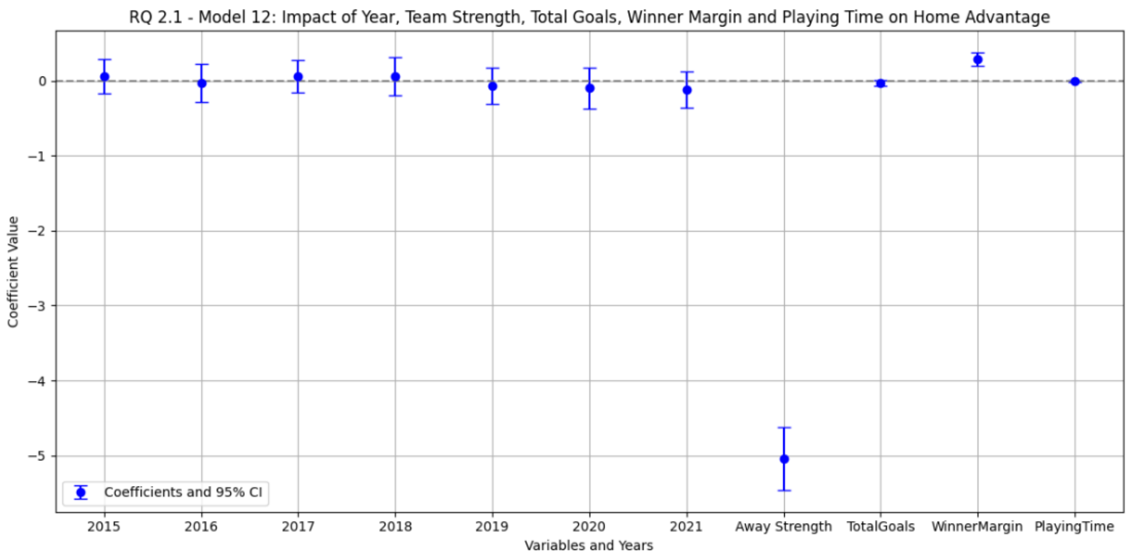

Figure 4.12. RQ 2.1 - Event Study (Model 12)

In model 12 (4.5), the coefficients for the individual years (2015-2021) are not statistically significant, confirming that after controlling for other variables in the model, there is no strong evidence of year-to-year variation in HA.

Overall, while there are no significant year-to-year changes in HA, the introduction of VAR is consistently associated with a slight decrease in HA. The strength of the away team and the margin of victory continue to be significant factors influencing HA. The inclusion of year effects does not dramatically change the relationship between the other variables and home advantage, indicating robust findings across different model specifications.

$$\begin{aligned} HomeAdvantage_i &= \beta_0 + \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\ &+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 PlayingTime_i \\ &+ \beta_5 Treatment_i \times PlayingTime_i + \beta_6 TotalGoals_i \\ &+ \beta_7 Treatment_i \times TotalGoals_i + \beta_8 WinnerMargin_i \\ &+ \beta_9 Treatment_i \times WinnerMargin_i + e_i \end{aligned} \tag{4.6}$$

Model 13 suggests that the direct impact of VAR on home advantage is not statistically significant when considering interaction effects with other variables. However, the strength of the away team and the margin of victory continue to be significant factors. The lack of significance in interaction terms indicates that the relationship between VAR introduction and other factors does not significantly change pre and post-VAR in terms of affecting home advantage. The model explains a similar proportion of the variability in home advantage as the previous models (30.5%).

4.3.2.2. Home Advantage as Home Wins

The model 1 of the framework yield a positive coefficient for VAR Introduction but not statistically significant. The introduction of VAR does not show a significant impact on the likelihood of home wins in either of the 6 models. This consistency across models suggests that VAR's implementation may not have fundamentally altered the dynamics of home team victories. A logit approach was assessed based on Home Wins being a binary variable, as already seen.

$$\log\frac{1-P(HomeWin)}{P(HomeWin)} = \beta_0 + \beta_1 Treatment_i + \beta_2 TeamStrenght_i + \beta_3 TotalGoals_i + \beta_4 PlayingTime_i + \beta_5 WinnerMargin_i + \varepsilon_i \tag{4.7}$$

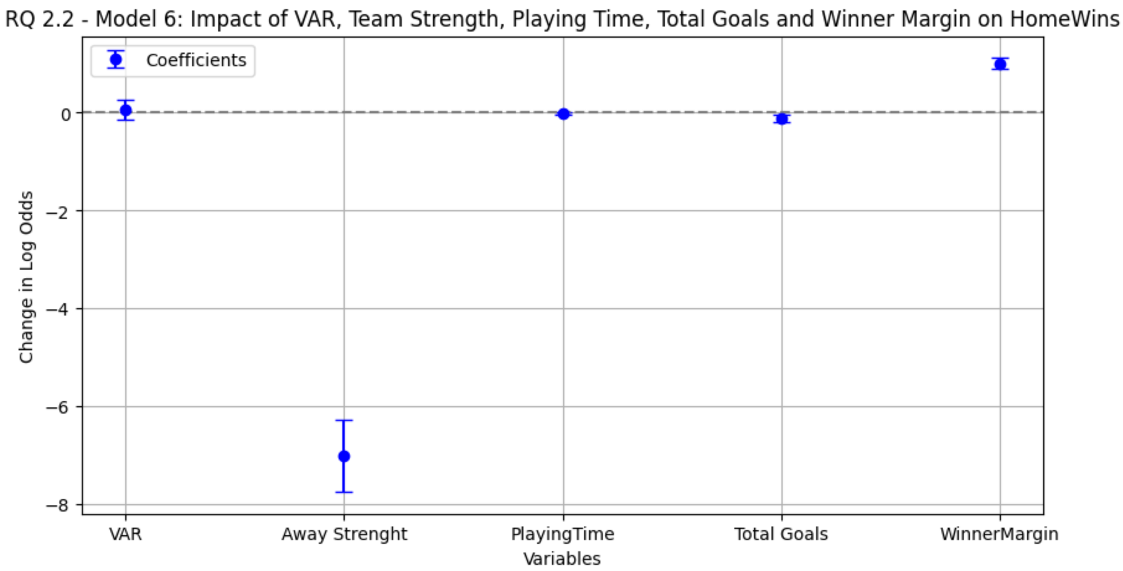

Figure 4.13. RQ 2.2 - Event Study (Model 6)

The strength of the away team and the margin of victory are significant predictors of home wins, with stronger away teams decreasing and larger winning margins increasing the likelihood of home wins in model 6 (4.7). The total number of goals and playing time are also significant factors, indicating their influence on the outcomes of home team victories. VAR shows no significant changes. Model 6 was tested for autocorrelation with a score of 1.97, which suggest that there is little to no autocorrelation in the residuals.

$$\log\frac{1-P(HomeWin_{it})}{P(HomeWin_{it})} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 TotalGoals_{it} + \beta_3 PlayingTime_{it} + \beta_4 WinnerMargin_{it} + \sum_{t\,=\,2014}^{2021} yi\ Year\ t\ + \varepsilon \tag{4.8}$$

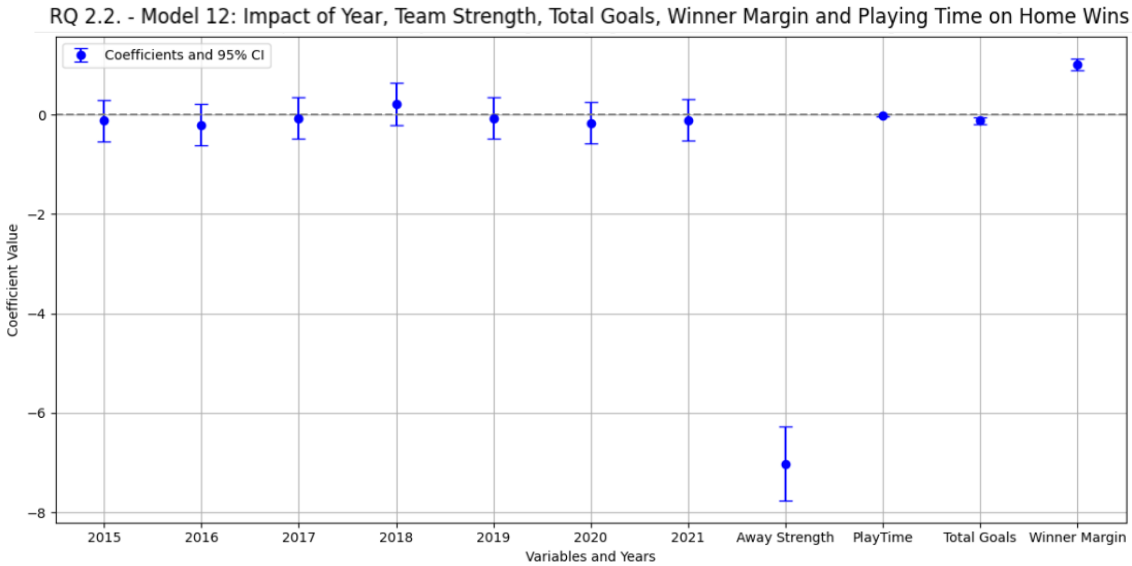

Figure 4.14. RQ 2.2 - Event Study (Model 12)

Yearly indicators from 2015 to 2021 in model 12 (4.8) do not significantly impact the likelihood of home wins, suggesting that factors other than time-specific events or trends are more influential.

Away team Strength emerges as a dominant predictor with a strong negative relationship, indicating that stronger away teams significantly reduce the likelihood of home wins, which is expected. Winner Margin, in contrast, continues to show a significant positive impact, reinforcing its importance in home victories.

$$\begin{aligned} HomeWin_i = \beta_0 &+ \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\ &+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 PlayingTime_i \\ &+ \beta_5 Treatment_i \times PlayingTime_i + \beta_6 TotalGoals_i \\ &+ \beta_7 Treatment_i \times TotalGoals_i + \beta_8 WinnerMargin_i \\ &+ \beta_9 Treatment_i \times WinnerMargin_i + e_i \end{aligned} \tag{4.9}$$

Model 13 (4.9) indicates that the direct impact of VAR on the likelihood of home wins is not significant when considering interaction effects with other variables. None of the interaction terms is statistically significant. This indicates that the impact of VAR on these variables does not differ notably before and after its introduction in terms of influencing home wins.

The interaction term between Treatment and Winner Margin is negative and marginally significant ($p < 0.1$), hinting that the introduction of VAR might slightly influence the relationship between the margin of victory and the likelihood of a home win.

The inclusion of interaction terms has not significantly altered the relationships observed in previous models. The main factors influencing home wins are consistent, with the introduction of VAR not showing a clear impact on HA.

4.3.2.3. Home Advantage as Home Points

Model 1 yield a negative coefficient for treatment despite being not statistically significant.

$$HomePoints_i = \beta_0 + \beta_1 Treatment_i + +\beta_2 TeamStrenght_i + \beta_3 TotalGoals_i + \beta_4 PlayingTime_i + \beta_5 WinnerMargin_i + \varepsilon_i \quad (4.10)$$

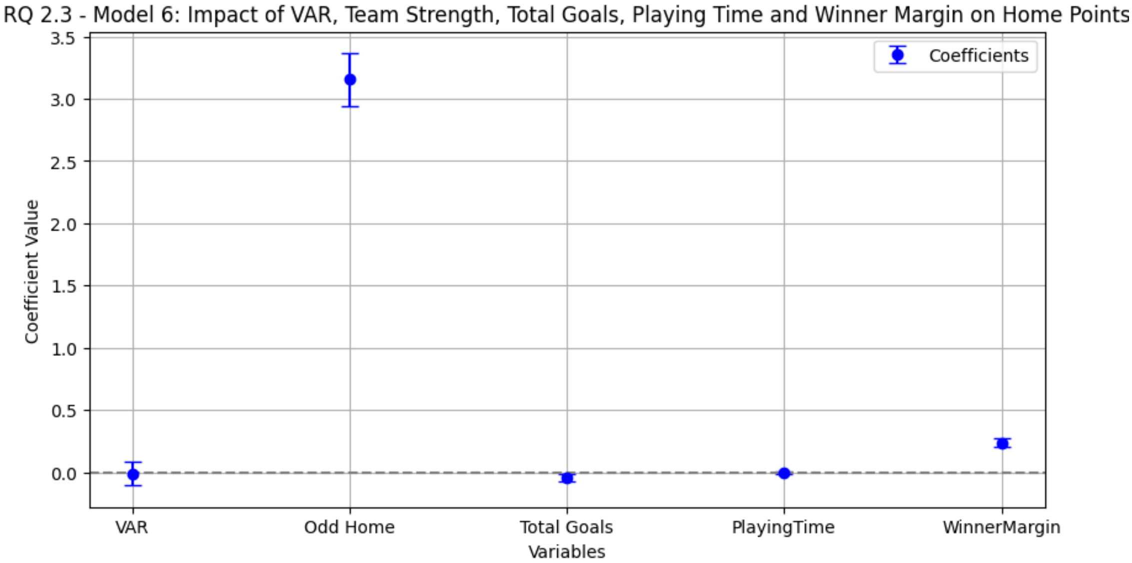

Figure 4.15. RQ 2.3 - Event Study (Model 6)

Similar to Model 1, the coefficient for treatment in Model 6 (4.10) is negative but not statistically significant, reinforcing the finding that VAR's introduction does not significantly alter the average points earned by the home team. As previously, there is no concerns regarding autocorrelation.

Home team Strength coefficient indicates that as the odds of the home team winning increase, so does the average points earn by the home team, which makes sense. The significant positive coefficient for Winner Margin suggests that larger margins of victory are associated with an increased average of points earned by the home team. In contrast, total goals have a negative but small significant impact on home points.

$$HomePoints_{it} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 TotalGoals_{it} + \beta_3 WinnerMargin_{it} + \beta_4 PlayingTime_{it} + \sum_{t=2014}^{2021} yi \;\; Year\, t \; + \varepsilon \tag{4.11}$$

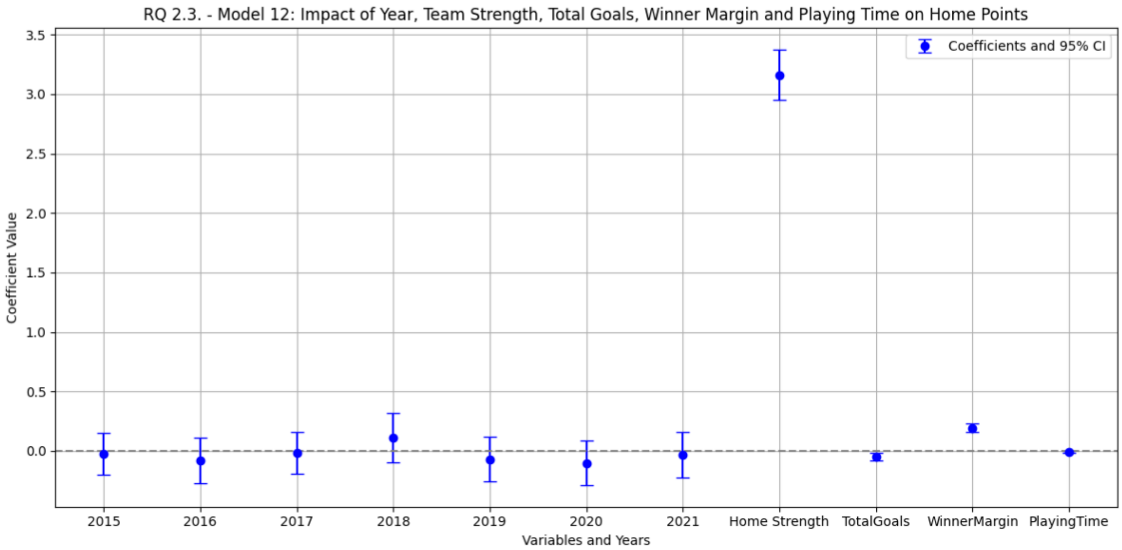

Figure 4.16. RQ 2.3 - Event Study (Model 12)

In model 12 (4.11), the coefficients for individual years from 2015 to 2021 show no statistically significancy. This suggests that after controlling for other variables in the model, there is no consistent year-to-year variation in home points that is statistically significant.

This model with year fixed effects suggests that while there are no significant year-to-year changes in home points, the odds of the home team winning and the margin of victory are significant factors in predicting the average points earned, while total goals slightly affect it too.

The VAR Introduction coefficient is negative but not statistically significant, indicating once more that VAR does not have a statistically significant direct impact on home points.

$$\begin{aligned} HomeAdvantage_i &= \beta_0 + \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\ &+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 PlayingTime_i \\ &+ \beta_5 Treatment_i \times PlayingTime_i + \beta_6 TotalGoals_i \\ &+ \beta_7 Treatment_i \times TotalGoals_i + \beta_8 WinnerMargin_i \\ &+ \beta_9 Treatment_i \times WinnerMargin_i + e_i \end{aligned} \tag{4.12}$$

Model 13 (4.12) coefficient for treatment is negative but also not statistically significant and none of the interaction terms are statistically significant. This suggests that the effect of VAR on the relationship between home points and variables like home odd to win, Playing Time, Total Goals, and Winner Margin is not significantly different pre and post-VAR introduction.

The influence of home odd to win and winner margin on home points remains consistent across all models, reinforcing their importance in predicting home points. However, in all models, including those with year fixed effects and interaction terms, the introduction of VAR does not have a significant direct impact on home points.

4.3.2.4. Home Advantage as Home Odds

In the following approaches, HA is assessed by the Home Odd to win the game, expressed as the probability of home to win the game. Home Advantage was chosen instead of Home Win and Home Points once it added more explanatory power to the models.

Model 1 yield a positive coefficient for treatment despite being not statistically significant.

$$Homeodd_i = \beta_0 + \beta_1 Treatment_i + +\beta_2 HomeAdvantage_i + \beta_3 TotalGoals_i + \beta_4 PlayingTime_i + \beta_5 WinnerMargin_i + \varepsilon_i \quad (4.13)$$

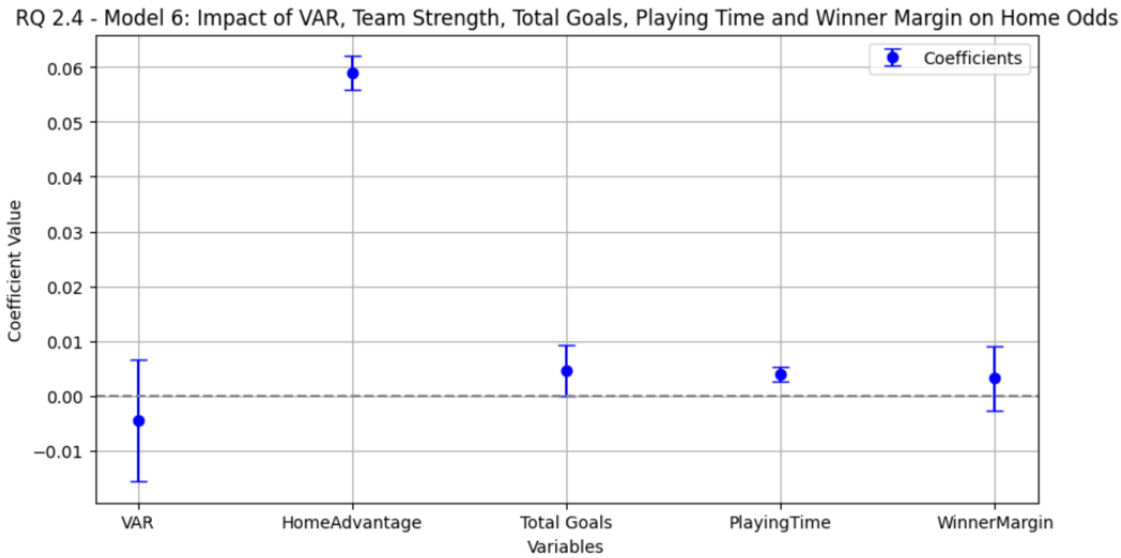

Figure 4.17. RQ 2.4 - Event Study (Model 6)

The R^2 value of model 6 (4.13) indicates that approximately 31.18% of the variability in the home team's winning probability is explained by the model. Model 6 suggests that the HA has a significant and positive impact on the probability of the home team winning, which means that betting odds are accurate as expected. Moreover, longer playing times also seem to favour the home team, while the total number of goals has a marginal positive effect. The introduction of VAR does not appear to significantly influence the odds of the home team winning.

This model was tested for autocorrelation, which yielded safe results.

$$Homeodd_{it} = \beta_0 + \beta_1 HomeAdvantage_{it} + \beta_2 TotalGoals_{it} + \beta_3 WinnerMargin_{it} + \beta_4 PlayingTime_{it} + \sum_{t=2014}^{2021} yi \ Year\ t + \varepsilon \quad (4.14)$$

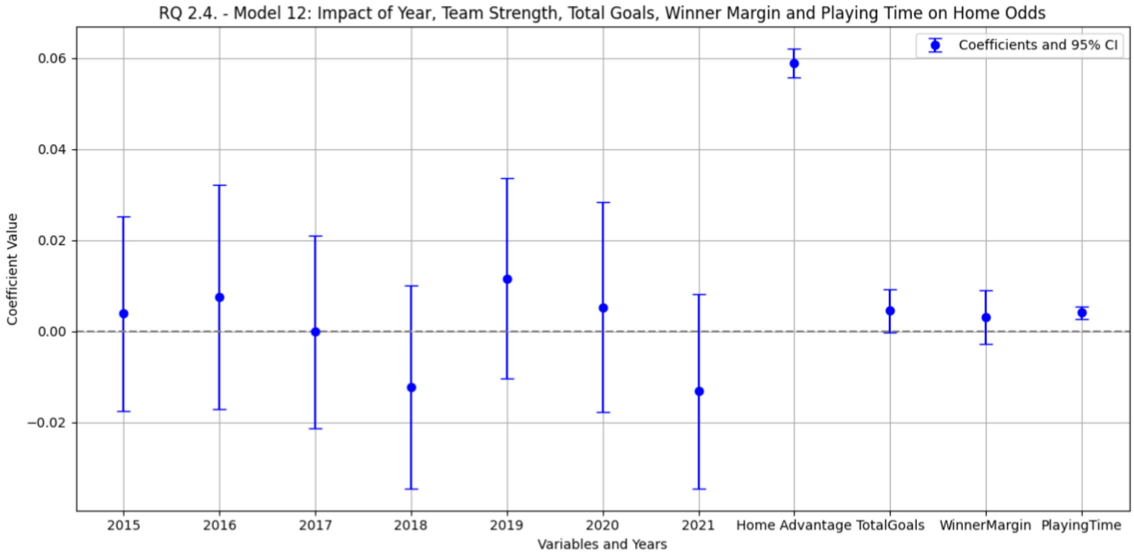

Figure 4.18. RQ 2.4 - Event Study (Model 12)

Model 12 (4.14) with year fixed effects underscores the importance of home advantage and playing time in influencing the home team's probability of winning. The yearly variations indicate that the likelihood of the home team winning has not been consistent across years, suggesting other contextual factors at play in different seasons.

$$\begin{aligned} HomeOdd_i = \beta_0 &+ \beta_1 Treatment_i + \beta_2 HomeAdvantage_i \\ &+ \beta_3 Treatment_i \times HomeAdvantage_i + \beta_4 PlayingTime_i \\ &+ \beta_5 Treatment_i \times PlayingTime_i + \beta_6 TotalGoals_i \\ &+ \beta_7 Treatment_i \times TotalGoals_i + \beta_8 WinnerMargin_i \\ &+ \beta_9 Treatment_i \times WinnerMargin_i + e_i \end{aligned} \quad (4.15)$$

Model 13 (4.15) indicates that home advantage and playing time continue to be significant factors in predicting the probability of a home team's victory. The introduction of VAR does not significantly impact this probability, nor does it significantly alter the relationships between the probability of a home win and the other variables considered in the model.

Overall, the 4 metrics used to assess the VAR impact in Home Advantage are consistent in showing that the direct impact of VAR introduction is not clear, with no significant direct effect found in most model specifications.

4.3.3. RQ 3 - VAR Impact on Playing Time

4.3.3.1. VAR Impact on Playing Time

The model 1 yield a highly significant treatment coefficient of 2.69, which suggests a statistically significant difference in the average playing time post-VAR introduction. However, this simple approach explains very little of the total variation in Playing Time (R^2 of 7.5%).

$$PlayingTime_i = \beta_0 + \beta_1 Treatment_i + +\beta_2 TeamStrenght_i + \beta_3 TotalGoals_i + \beta_4 HomeAdvantage_i + \beta_5 WinnerMargin_i + \varepsilon_i \tag{4.16}$$

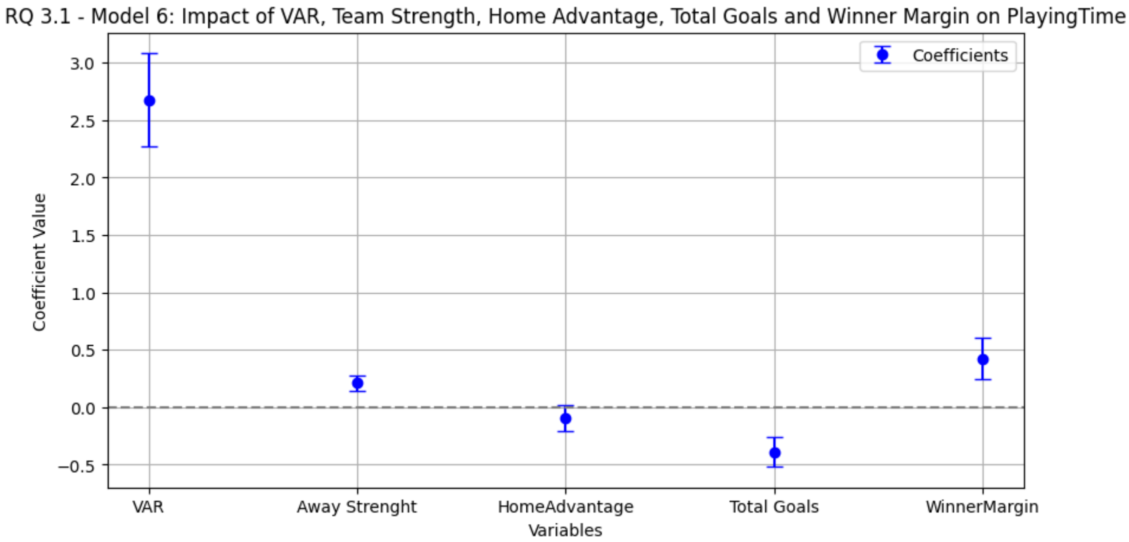

Figure 4.19. RQ 3.1 - Event Study (Model 6)

Regarding Model 6 (4.16) regression on Playing Time, it was away team's odd to win that better contributed to the model rather than home team odd. Additionally, home advantage variable expressed as goal difference was the better variable to fit the model regarding home advantage variables. The model explains 10.58% of the variance in Playing Time, as indicated by the respective adjusted R^2. As for the previous analysis, the Durbin-Watson test yielded safe results.

The VAR coefficient is 2.68 and is highly statistically significant in model 6 (4.16). This implies that there is a statistically significant increase of around 2 minutes and 40 seconds in the average playing time between the pre and post-VAR.

Away team strength is also positive and significant, which suggests that stronger away teams promote more playing time, which makes sense once time-wasting is a way for weak teams to seek wins against stronger teams (Greve et al., 2019). The HA coefficient is negative but only slightly significant (p < 0.1).

Total Goals and Winner Margin both contribute significantly to playing time, but in opposite directions, which is interesting. It might suggest that playing time is bigger when games are already solved (large difference between teams), but it decreases if there are a lot of goals, possibly games with several goals and with a close score.

$$
\begin{aligned}
PlayingTime_{it} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 TotalGoals_{it} + \beta_3 WinnerMargin_{it} \\
+ \beta_4 HomeAdvantage_{it} + \sum_{t\,=\,2014}^{2021} yi\ \ Year\ t\ + \varepsilon
\end{aligned}
\tag{4.17}
$$

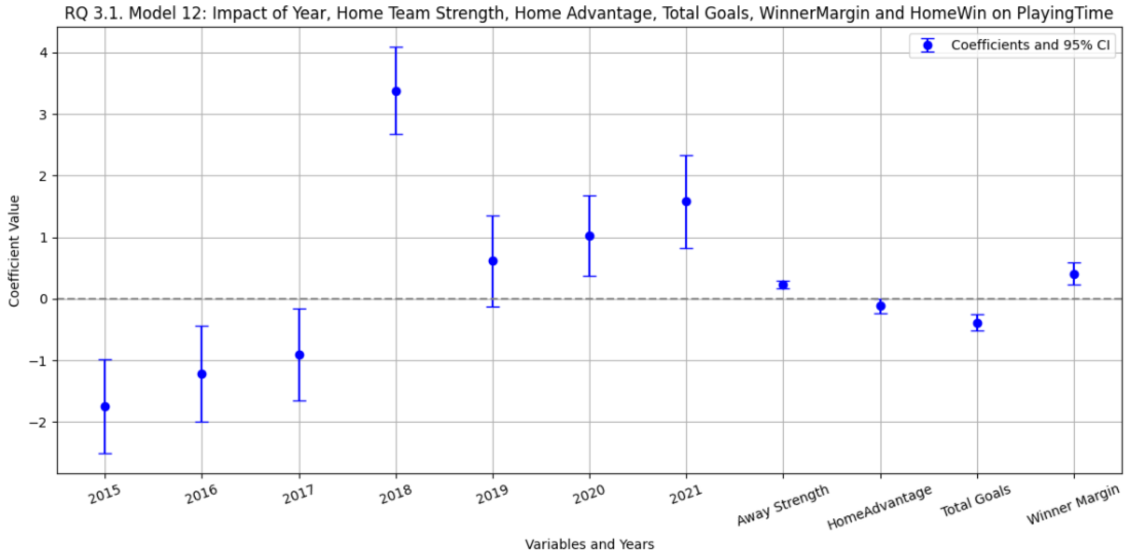

Figure 4.20. RQ 3.1 - Event Study (Model 12)

In model 12 (4.17), the coefficients for years, particularly for 2018 (when VAR was introduced), show a significant increase in playing time compared to the base year (2014). This indicates a distinct shift in playing time due to the introduction of VAR, alongside yearly variations.

The coefficients of Away Strength, Winner Margin and Total Goals remain consistent and statistically significant throughout all the models studied. HA seem to don't have a clear impact on total Playing Time.

$$
\begin{aligned}
PlayingTime_i = \beta_0 + \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\
+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 HomeAdvantage_i \\
+ \beta_5 Treatment_i \times HomeAdvantage_i + \beta_6 TotalGoals_i \\
+ \beta_7 Treatment_i \times TotalGoals_i + \beta_8 WinnerMargin_i \\
+ \beta_9 Treatment_i \times WinnerMargin_i + e_i
\end{aligned}
\tag{4.18}
$$

This final model 13 (4.18) indicates a strong and significant effect of VAR on increasing playing time in football matches. This effect seems to be modulated by the away team strength, with the impact of VAR being less pronounced in matches where the away team is weaker. The interaction with other variables like home advantage, total goals, and winner margin is not significant, suggesting that the primary effect of VAR on playing time is consistent across various match conditions.

The negative and significant coefficient (-0.1928 and $p < 0.05$) between VAR and Away team strength is quite interesting. One possible conclusion is that while VAR generally leads to longer games, the extent of this increase in playing time is influenced by the competitive balance between the home and away teams, being slightly less in games where the away team is much weaker, which robust previous conclusions.

In summary, the introduction of VAR has a clear and quantifiable impact on extending the playing time of football matches.

4.3.3.2. VAR Impact on Effective Playing Time

In coherence with playing time, the model 1 yield a highly significant treatment coefficient, despite its low size.

$$\begin{aligned} &EffectivePlayingTime_i \\ &\qquad = \beta_0 + \beta_1 Treatment_i + +\beta_2 TeamStrenght_i + \beta_3 TotalGoals_i \\ &\qquad + \beta_4 HomeAdvantage_i + \beta_5 WinnerMargin_i + \varepsilon_i \end{aligned} \tag{4.19}$$

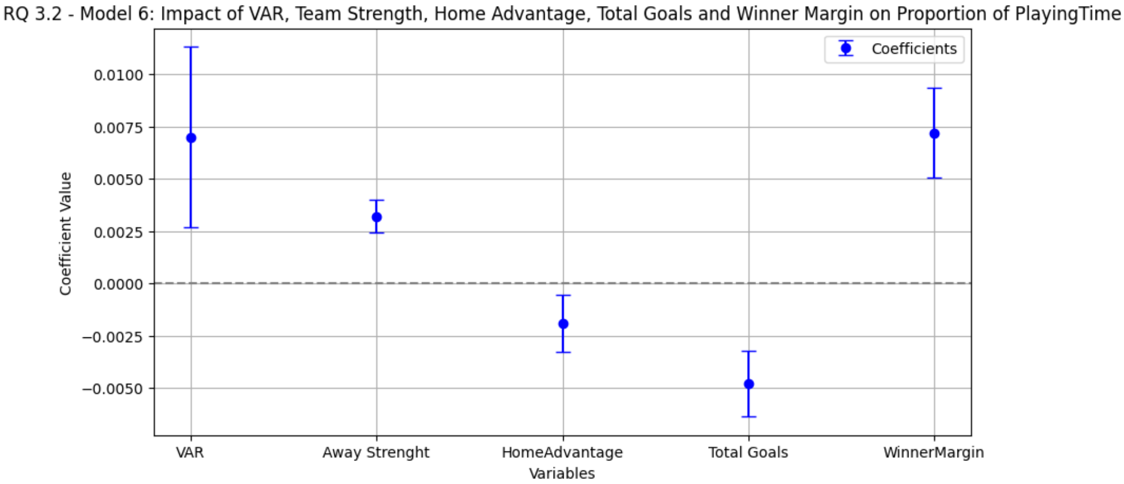

Figure 4.21. RQ 3.2 - Event Study (Model 6)

In the same way of previous analysis, safe results were achieved for autocorrelation in model 6 (4.19), which states that all variables are significantly impacting the effective playing time. The VAR introduction seems to increase the proportion of actual playing time, suggesting a potential efficiency gain in match management. Even though the coefficient accounts for an increase of only 0.7%, this would mean around 38 seconds on a 90-minute game, which, despite not seem like much, in the fast-paced environment can be crucial.

$$EffectivePlayingTime_{it} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 TotalGoals_{it} + \beta_3 WinnerMargin_{it} + \beta_4 HomeAdvantage_{it} + \sum_{t=2014}^{2021} yi \; Year\, t \; + \varepsilon \tag{4.20}$$

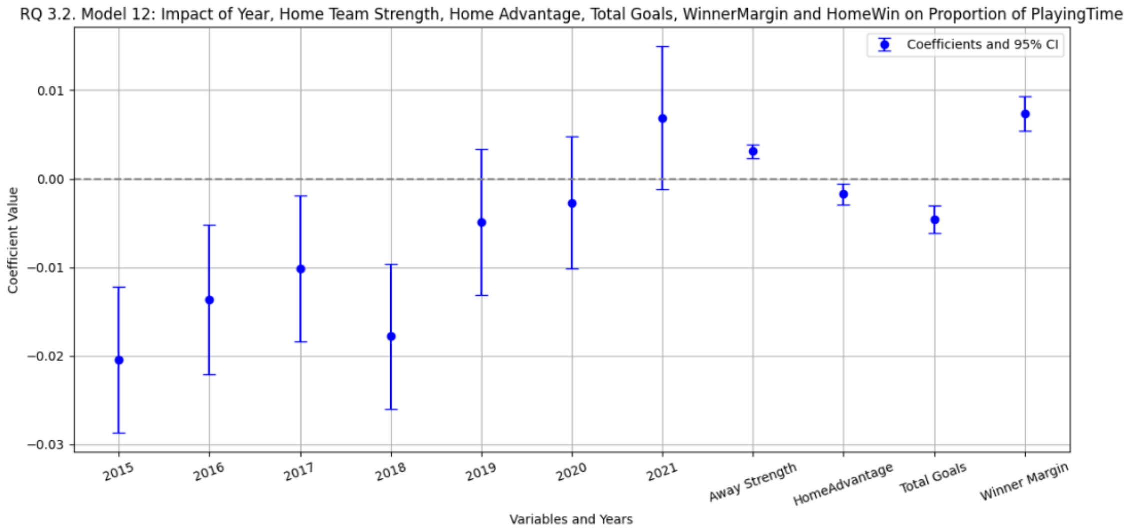

Figure 4.22. RQ 3.2 - Event Study (Model 12)

The model 12 with year fixed effects (4.20) suggests that while there are some variations in the effective playing time across different years and game conditions, the introduction of VAR doesn't seem to have a significant direct impact on this metric, which contrasts with previous analysis without year fixed effects.

The year fixed effects model captures specific yearly variations that could mask or enhance the perceived impact of VAR when not accounted for. By including year fixed effects, the direct impact of VAR is more isolated. This means the model is better at distinguishing the effect of VAR from other concurrent changes over the years.

$$\begin{aligned} Effective\ PlayingTime_i &= \beta_0 + \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\ &+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 HomeAdvantage_i \\ &+ \beta_5 Treatment_i \times HomeAdvantage_i + \beta_6 TotalGoals_i \\ &+ \beta_7 Treatment_i \times TotalGoals_i + \beta_8 WinnerMargin_i \\ &+ \beta_9 Treatment_i \times WinnerMargin_i + e_i \end{aligned} \tag{4.21}$$

In model 13 (4.21), the coefficient for treatment (0.0122, $p < 0.05$) indicates a statistically significant increase in the percentage of playing time following the introduction of VAR.

The interaction terms between VAR introduction and other variables are not significant, which indicates that a possible increase in the effective playing time is not affected by other variables.

These varying results highlight the complexity of measuring VAR's impact and how different model specifications can lead to different interpretations. The simple model suggests a clear impact, the fixed year effects model suggests no clear impact, and the interaction terms model reinforces the initial findings of the simple model but with a more nuanced understanding.

4.3.4. RQ 4 - VAR Impact on Competitiveness

In the process of modelling the impact of VAR on the competitiveness of matches, the variable Playing Time was excluded as a control variable. This decision was grounded on the fact that its inclusion provided less explanatory value and a negative intercept was observed in the model, raising concerns about the interpretability and intuitive understanding of the model's results.

Sensitivity analyses revealed that the key findings regarding the impact of VAR on Winner Margin were robust to the exclusion of Playing Time. This consistency in results supports the decision to prioritize a more parsimonious model that directly addresses the research question without the inclusion of additional control variables that do not substantially alter the study's conclusions.

In this sense, model 1, accounting simple for VAR Introduction impact on Winner Margin, provided a non-significant impact of VAR in competitiveness of games.

$$\begin{aligned} WinnerMargin_i \\ &= \beta_0 + \beta_1 Treatment_i + +\beta_2 TeamStrenght_i + \beta_3 TotalGoals_i \\ &+ \beta_4 HomeAdvantage_i + \varepsilon_i \end{aligned} \tag{4.22}$$

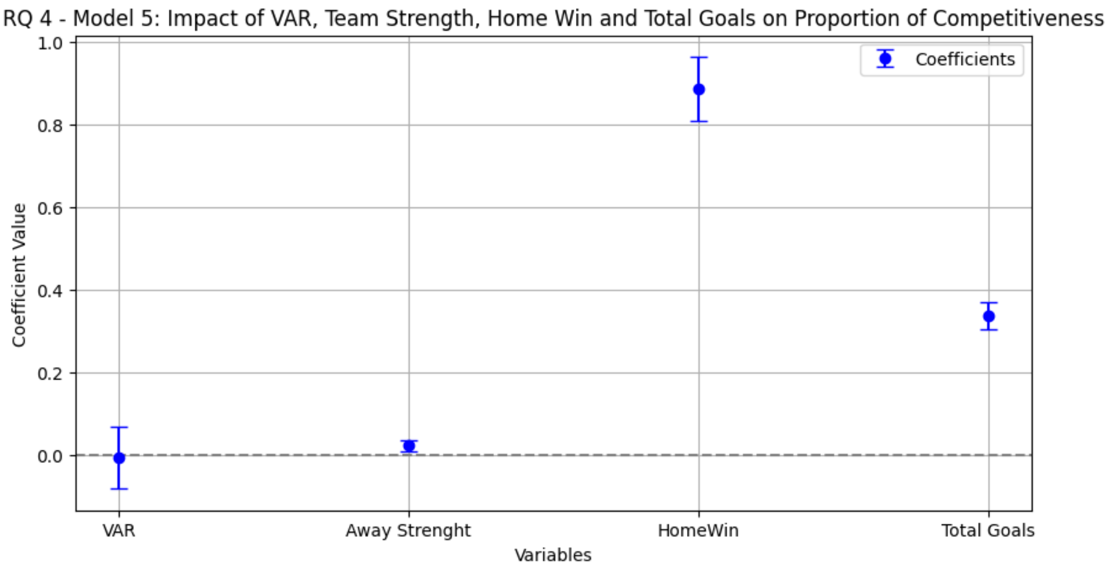

Figure 4.23. RQ 4 - Event Study (Model 5)

In conclusion, this model 5 (4.22) suggest that while the introduction of VAR does not significantly affect the Winner Margin in football matches, other factors like whether the home team wins, the odds for the away team, and the total number of goals significantly influence the margin by which teams win or lose.

The variable removal did not alter the safe results obtain when testing for autocorrelation.

$$WinnerMargin_{it} = \beta_0 + \beta_1 TeamStrenght_{it} + \beta_2 TotalGoals_{it} + \beta_3 HomeAdvantage_{it} + \sum_{t=2014}^{2021} yi\ Year\ t\ + \varepsilon \quad (4.23)$$

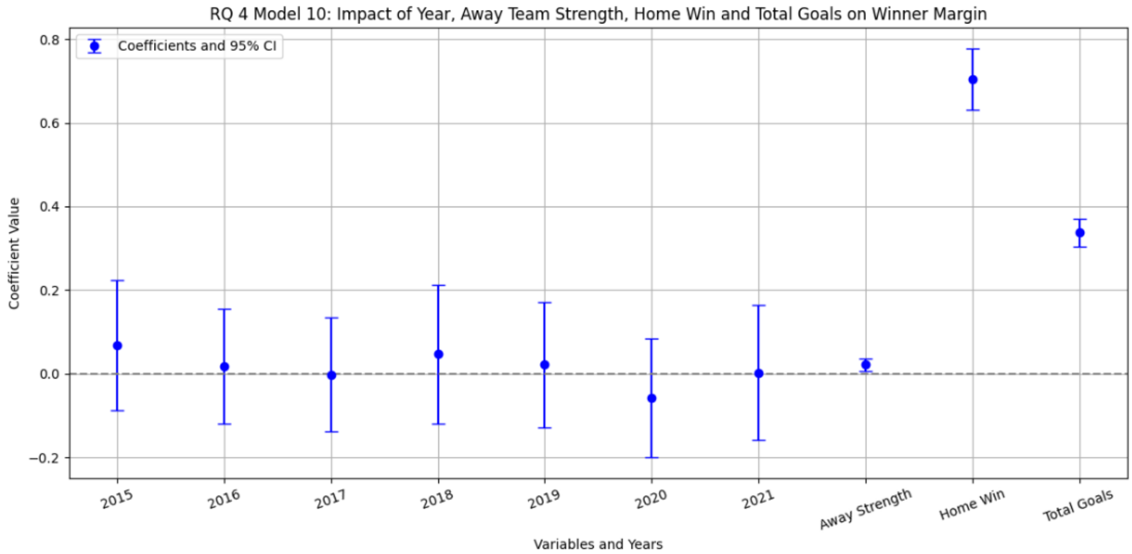

Figure 4.24. RQ 4 - Event Study (Model 10)

In model 10 (4.23), the coefficients for individual years, compared to the reference year (2014), show varied impacts on the Winner Margin and none is statistically significant. This indicates that the introduction of VAR may have had a marginal impact on the winner goal difference in games, but this effect is relatively small. The control variables coefficients are statistically significant, which remains consistent with the previous model 5.

$$\begin{aligned} WinnerMargin_i &= \beta_0 + \beta_1 Treatment_i + \beta_2 TeamStrenght_i \\ &+ \beta_3 Treatment_i \times TeamStrength_i + \beta_4 HomeAdvantage_i \\ &+ \beta_5 Treatment_i \times HomeAdvantage_i + \beta_6 TotalGoals_i \\ &+ \beta_7 Treatment_i \times TotalGoals_i + e_i \end{aligned} \quad (4.24)$$

Adding interaction terms (4.24) robust the idea that Home Wins, Away Strenght and Total Goals are significant predictors of how close a game is, decreasing the competitiveness of a game, if measured as the goal difference between the winner and looser.

The coefficient for VAR introduction continued to be not statistically significant, implying that the introduction of VAR doesn't have a direct significant impact on the Winner Margin.

The negative and significant coefficient on the interaction between VAR and Home Win suggests that the introduction of VAR might reduce the Winner Margin in matches where the home team wins. The remaining interaction terms are not statistically significant, indicating that VAR introduction does not significantly alter the relationship between these variables and the Winner Margin.

The analysis indicates that VAR introduction has not notably changed the competitiveness of games in terms of goal difference.

5. Robustness Checks

The robustness checks account for both GLM and RDD performed for all of the research questions. Regarding GLM, the results are presented in this chapter and the respective regression output is shown in the Appendix C - Generalized Linear Models. For RDD, this chapter presents visualization for each analysis and the respective regression output is presented throughout the Appendix D - Regression Discontinuity Design.

5.1. RQ 1- VAR Impact on Goals

Generalized Linear Models

Since data on total goals is count data, cannot handle negative values and is not over dispersed (mean and variance are slightly different), the GLM assumed a Poisson distribution.

Comparing to the Event Studies regressions, GLM confirms the suggestion that the introduction of the VAR is associated with a slight increase in the total number of goals scored, being that the effect is statistically significant. Moreover, winner margin and playing time remain statistically significant with opposite directions in its impact.

When considering fixed effects, the output suggests that there is a year effect on the total goals, with 2018 being the only year showing positive and statistically significant coefficient. Playing time and Winner Margin kept its significancy.

Regression Discontinuity Design

Both results with a bandwidth of 2 and 3 years suggest a shift in total goals after VAR Introduction. The RDD graph shows a discontinuity at the cut-off point (introduction of VAR in 2018), which suggest that the introduction of VAR has a significant effect on the total goals scored in matches (p value 0.011).

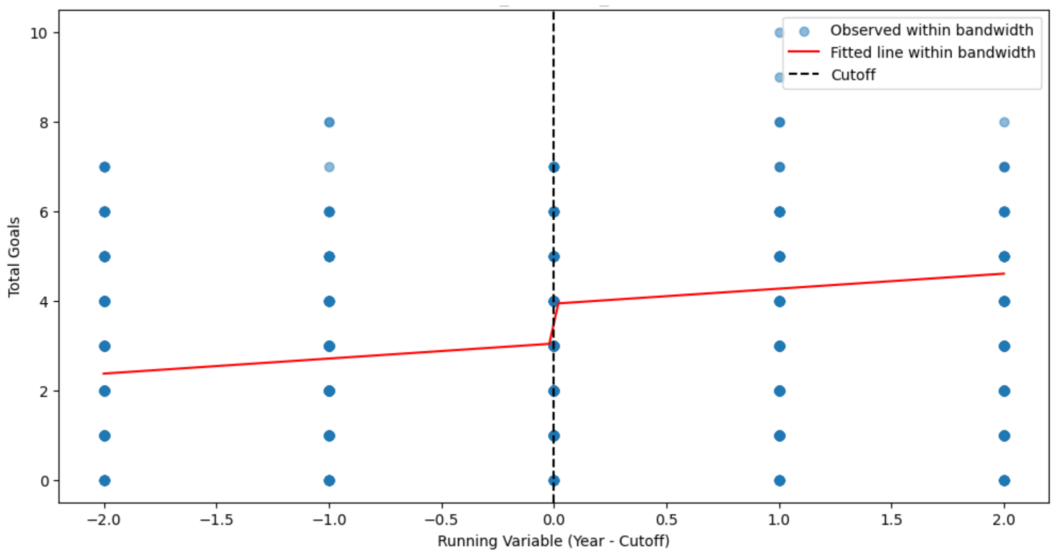

Figure 25.1. RQ 1 – Total Goals: RDD w/bandwidth = 2

5.2. RQ 2- VAR Impact on HA

5.2.1. Home Advantage as Home Goal Difference

Generalized Linear Models

As previously seen in the Descriptive Statistics, the variable HA accounts for negative values, which mean that the home team lost the game and, therefore, the distribution cannot be considered as a Poisson distribution. As such, a gaussian distribution will be assumed despite the mean and the variance being considerable different, which indicates overdispersion and awards caution on conclusions.

GLM did not yield significative values for the VAR introduction and, when accounting for fixed effects, the results are consistent: the coefficients for the individual years are not statistically significant, and away strength and winner margin remain significant.

Regression Discontinuity Design

This visual representation is consistent with the previous conclusions that the introduction of VAR does not show a significant change in the home advantage, once there is not visible cutoff in the continuity, which is sustained by a p value of 0.813.

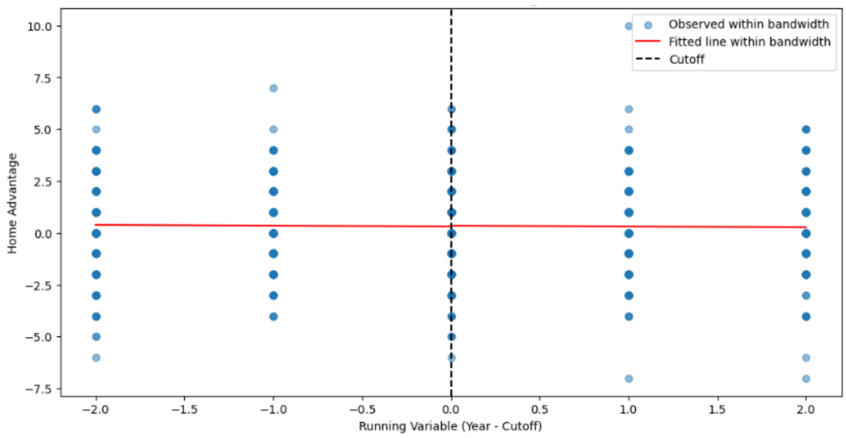

Figure 5.26. RQ 2.1 - Home Advantage: RDD w/bandwidth = 2

5.2.2. Home Advantage as Home Wins

Generalized Linear Models

Once GLM supports binary variables, the same test can be applied, with a slightly change of distribution family from Gaussian to Binomial.

In both GLM (with treatment variable and accounting for year effects), the effect of VAR introduction is not significant, indicating that its introduction hasn't markedly changed the likelihood of home wins. Moreover, the away strength and winner margin remain significant.

Regression Discontinuity Design

Due to variable specifications addressed in Materials and Methods, this particular RDD was regressed using a simple logistic regression and a second-order polynomial regression.

Both visualizations yield a positive spike around the cutoff, suggesting the idea of an increase of Home Wins after VAR introduction. Regarding the simple logistic regression (Figure 24), this spike is significant ($p < 0.05$) both for 2,3 and 4 years as bandwidth. The 2nd order logistic regression, despite also showing a positive spike, none of the several regressions yield a significant increase in Home Wins in post-VAR period.

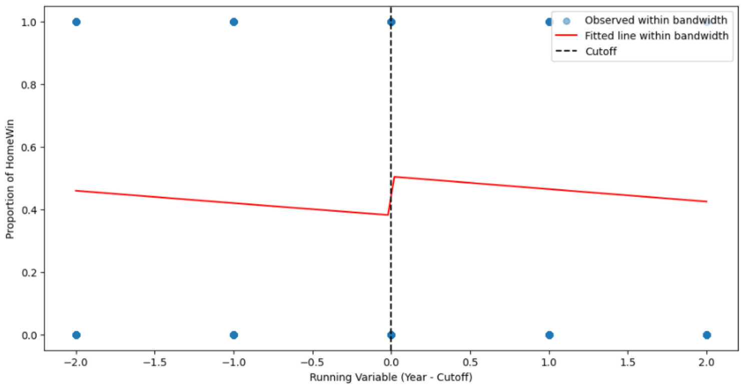

Figure 273. RQ 2.2 - Home Win: RDD 1st order Logistic Regression w/bandwidth = 2)

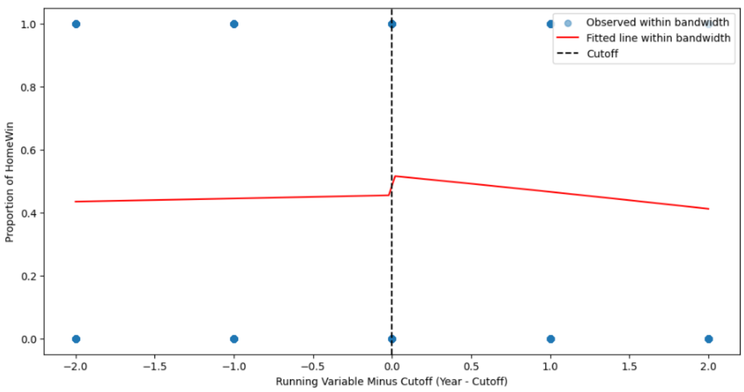

Figure 284. RQ 2.2 - Home Win: RDD 2nd order Logistic Regression w/bandwidth = 2

5.2.3. Home Advantage as Home Points

Generalized Linear Models

The variable Home Points is somehow over dispersed and therefore a Negative Binomial (NB) was used, once it deals better with count data that exhibits overdispersion. NB has an additional parameter alpha that controls the degree of overdispersion.

Both models of GLM robust earlier conclusions that VAR does not have a direct impact on the points gained by home teams. Moreover, this models also confirm the significance of total goals, winner margin and home strength when prediction home points.

Regression Discontinuity Design

The jump at the cutoff point, which is not very pronounced, is meant to represent the effect of the VAR Introduction. However, the regression output suggests that this effect is not statistically significant (p value of 0.113).

The lack of a substantial discontinuity at the cutoff in the graph and the non-significance of the treatment effect in the regression output imply that there is not enough evidence to conclude a strong impact of VAR on Home Points in the immediate term around its introduction.

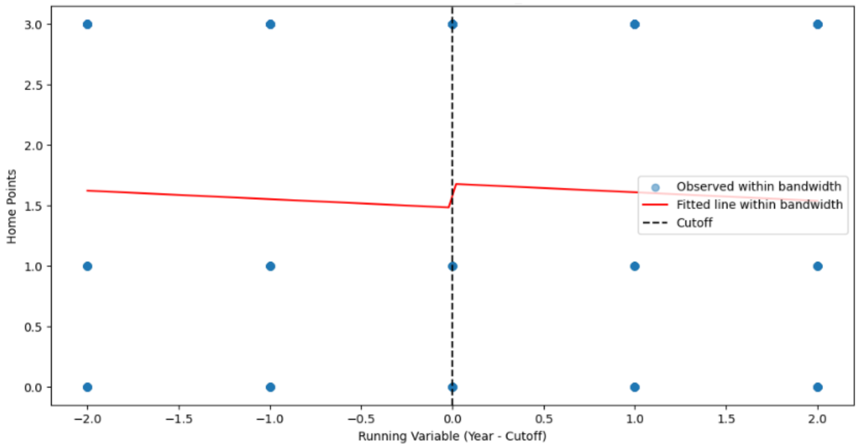

Figure 29. RQ 2.3 – Home Points: RDD w/bandwidth = 2

5.2.4. Home Advantage as Home Odds

Generalized Linear Models

GLM used a gaussian distribution to assess VAR Impact on Home Odds. All the control variables kept its direction of impact on Home Odds and Home Advantage, Total Goals and Playing Time kept its significancy levels and VAR continued to show no significant effect on home odds.

When accounting for fixed year effects, GLM also confirms the idea of a slightly decrease on home's probability to win a game, despite not being statistically significant. The control variables also maintain their significancy levels.

Regression Discontinuity Design

The visual inspection of the RDD shows no significant spike on home odds regarding the VAR Introduction, which has a non-significant impact (p value 0.489), robust previous conclusions.

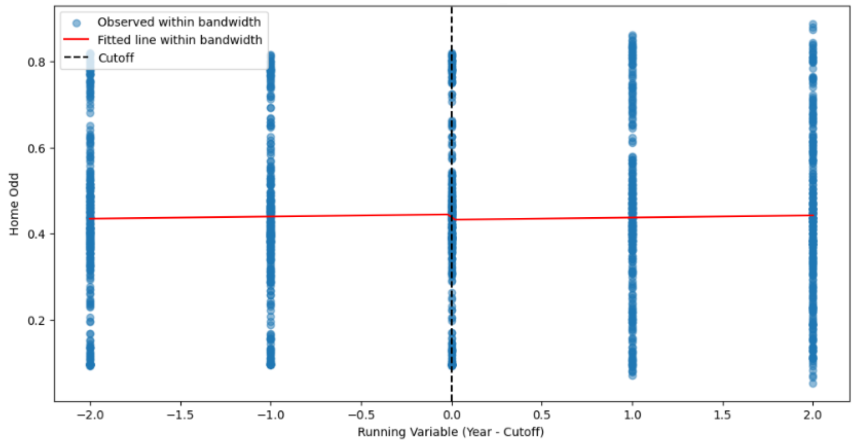

Figure 30. RQ 2.4 – Home Odd: RDD w/bandwidth = 2

5.3. RQ 3- VAR Impact on Playing Time

5.3.1. Playing Time

Generalized Linear Models

The GLM assumed a Gaussian distribution and confirmed the statistically significant positive coefficient for Treatment, Away Strength, Winner Margin. Moreover, it also robust the hypothesis that total goals decrease playing time.

The year conclusions are also backed by GLM, considering the coefficient of 2018 highly significant ($p < 0.01$).

Regression Discontinuity Design

The RDD model also robust the conclusion that VAR Introduction indeed contributed to an increment on Playing Time in Portuguese Football League, as explained by the jump at the cutoff and the statistically significant coefficient ($p < 0.001$).

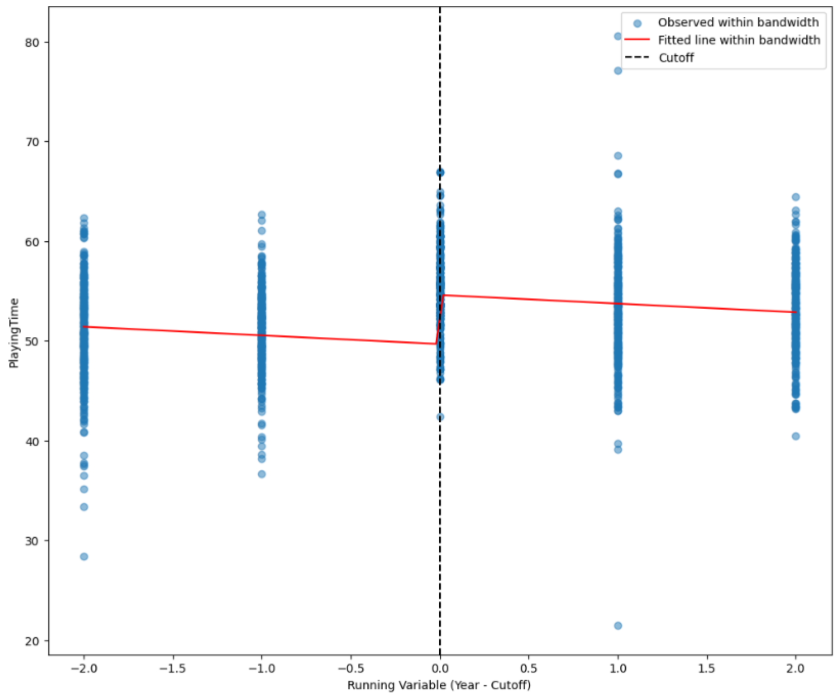

Figure 31. RQ 3.1 – Playing Time: RDD w/bandwidth = 2

5.3.2. Effective Playing Time

Generalized Linear Models

Using a gaussian GLM, the event study results are confirmed once again that VAR Introduction has highly statistically impact ($p < 0.001$).

The year conclusions are also stable affirming that the VAR introduction increased the effective of playing time in its first year.

Regression Discontinuity Design

The graphic cutoff highlights the statistically significant decrease in the effective playing time in VAR first year of use. The same regression, with a bandwidth of 4 years (instead of 2 years) yield a non-significant coefficient (p value = 0.204), which still does not agree with previous conclusions.

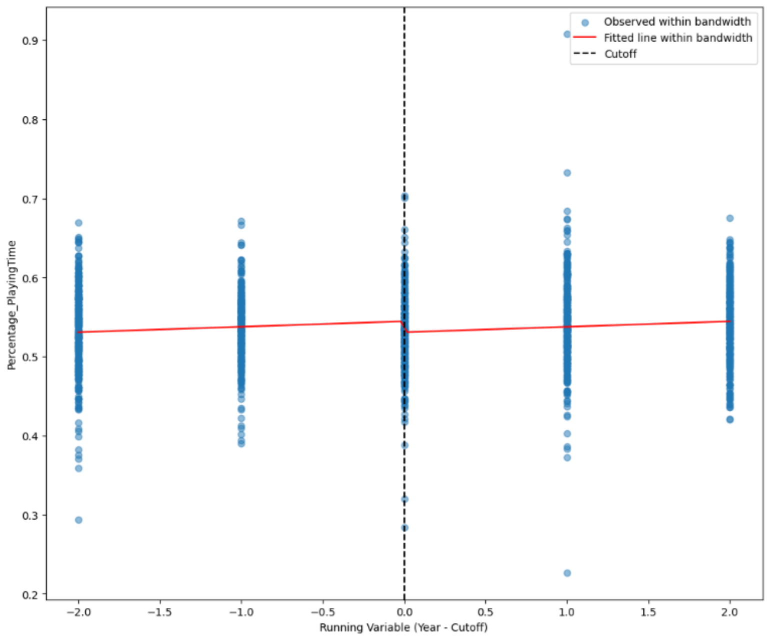

Figure 32. RQ 3.2 – Effective Playing Time: RDD w/bandwidth = 2

5.4. RQ 4- VAR Impact on Competitiveness

Generalized Linear Models

Considering that 24% of this variable is zeros, the GLM assumed a Zero Inflated Poisson. This approach was not made with Home Advantage (Goal difference between home and away) once it accounted for negative values.

The GLM assessed robust the previous studies results with a non-statistically significant coefficient for VAR impact on competitiveness, confirming home wins and total goals as increasing factors in this discrepancy, as expected. Regarding year analysis, GLM yield similar results with no year being significant.

Regression Discontinuity Design

The treatment effect is positive but not statistically significant (p=0.257), confirming that the introduction of VAR does not have a statistically significant effect on the Winner Margin at the conventional levels of significance.

The fitted line within the bandwidth does not appear to be a substantial jump or drop at the cutoff point. The absence of a sharp discontinuity suggests that the introduction of VAR did not lead to an immediate and significant change in the Winner Margin.

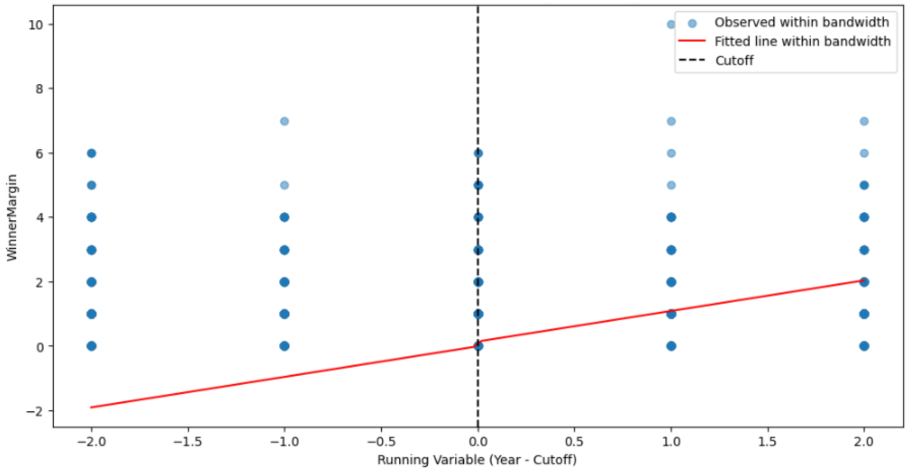

Figure 33. RQ 4 – Winner Margin: RDD w/bandwidth = 2

6. Discussion

6.1. RQ 1 – VAR Impact on Total Goals

Regarding the VAR Impact on total goals, descriptive statistics suggested an uptick in goals, particularly for away teams, although Mann Whitney U test results questioned the statistical significance of this increase.

Additionally, Event studies indicated a significant initial rise in total goals with the advent of VAR in 2018, but this effect seemed to decrease in the following years, which can be due to Covid-19 impact, as stated in other studies. However, introducing interaction terms into the model soften this clear impact, suggesting a more complex relationship between VAR and other game variables.

Home Wins and playing time are shown to significantly decrease total goals, which suggests that home wins tend to be 'cleaner' with less goals and that goals and playing time have a negative relationship, as also concluded in the playing time analysis. On the other side, Winner Margin, as expected, increase the total goals in a game and so does home team Strength.

The apparent contradiction between the effects of Home Wins and Home Strength can be elucidated by the Portuguese league's scoring patterns—where the majority of matches (54%) end with two or fewer goals, as perceived in boxplot presented in descriptive statistics. It implies that Home Wins are often narrow, not reflecting the decisive goal margins expected from stronger teams.

The GLM supported the notion of VAR's significant impact on total goals, also affected by the winning margin and playing time. RDD visualization further reinforced these findings, showing a clear spike at VAR's introduction point.

In summary, the coherence across various analytical methods underscores the significant influence of VAR on match scoring.

While direct comparisons with the Portuguese league are unavailable, this research aligns with Errekagorri et al. (2020), which suggests that VAR interventions correlate with a rise in total goals. In contrast, this study's results differ from Carlos et al. (2019), where VAR introduction in Serie A reportedly led to fewer goals, and from Gürler & Polat (2021), which observed a decline in goals in the Turkish league following VAR's implementation.

Besides being from a different country, these contrasting studies are limited by their scope, examining only one or two seasons before and after the adoption of VAR which is fairly insufficient to establish long-term trends or account for pre-existing dynamics. By extending the analysis over multiple seasons and employing comprehensive statistical techniques, this study provides a more nuanced understanding of VAR effects on the game.

The observation that the introduction of VAR contributes to an increase in goals can be a source of optimism for fans. It suggests that VAR, rather than diminishing the excitement of the game, enhances the thrill of goal celebrations and the overall enjoyment of football.

6.2. RQ 2- VAR Impact on Home Advantage

HA was measured with a multi analysis on 4 different metrics: Home Goal Difference, Home Wins, Home Points and Home Probability to win the game.

Goal difference between home teams and away teams (RQ 2.1) decreased 0.08 goals per game, mostly derived from away teams scoring more goals after VAR introduction. These reduction in HA, however, it not statistically significant when using a Mann Whitney U test.

All the event studies show a non-significant impact in home advantage measured by team's goal difference. Away strength shows a negative significant impact on HA, which makes sense, and winner margin have a positive significant impact, suggesting that larger winner margins are related to home wins, which also makes sense in the light of traditional HA.

GLM robust the idea of a non-significant impact of VAR and RDD show no spike in data, concordantly with previous conclusions.

When measured as Home Wins (RQ 2.2), HA shows a slight reduction when simply accounting for more away wins, which is softened with a non-significant Chi Square and Fisher's exact test.

Event study models agree that VAR Introduction has not change the extent of home wins, and that away strength and winner margin significant impact the home wins. Away strength naturally is expected to decrease HA and larger winner margins will tend to be in favour of home teams.

GLM yield a non-significant impact too, confirming the control variables impact. Interestingly, RDD showed a significant positive spike in the data, which suggests that VAR actually increase home wins. However, for a slightly more complex model with a 2^{nd} order logistic regression, these results become non-significant, confirming the previous results of no impact.

RQ 2.3, which measures the average gained points by the home team, showed that home teams win slightly less points at home, but not significant when Mann Whitney U test is assessed.

Event studies show no significant change in Home Advantage, yielding a positive impact for home strength and winner margin, being total goals negative. This contrast between larger winner margin and total goals is also interesting, meaning that total goals might be related to away goals, which could induce future research.

GLM and RDD both robust findings of non-significant VAR impact on Home Advantage when measuring it by home points.

When assessing HA based on home betting odds (2.4), no major changes were found with a Mann Whitney U test approach for HA. Event studies, confirm the simple idea of a non-significant impact in home odds, which states that the normalized probability of a home win is roughly the same. HA (goal difference) is a positive predictor of home probability of win, which is expected and playing time is also positively associated with bigger home's probability of win, which could promote interesting studies the fairness of increasing playing time of a game.

GLM and RDD both show a non-significant decrease in HA measured as home odds to win.

The multi approach of HA in this study leads to the conclusion that the VAR has likely not significantly altered HA. These findings contrast with observations from other leagues such as the Chinese Super League (Han et al., 2020) and both Bundesliga and La Liga (Holder et al., 2021), where a decrease in HA has been noted. However, and as previously referred, these studies are limited in their scope, covering only 2 to 4 seasons, and employing singular metrics to assess HA. In line with our results, the DiD approach employed by Kim et al. (2022) also found an increase in away goals and stable home goals.

6.3. RQ 3 – VAR impact on Playing Time

Regarding the hot topic of Playing Time (RQ 3.1) and Effective Playing Time (3.2), simply descriptive analysis shows an increase in both metrics and Mann Whitney U test confirm this significant increase, however smaller in effective playing time.

Event studies show a significant increase in playing time associated with VAR introduction, despite model 13 (interaction terms) showed no impact. Away strength and winner margin increase playing time, which yields distinct conclusions: (1) stronger away teams might make the game more active and increase playing time; (2) larger winner margins have less wasted time, once the result might not in dispute anymore (weaker teams waste more time when trying to win). However, these conclusions would require a more comprehensive study. In addition, games with more goals scored decrease playing time, which suggest a result of 3-2 would have less playing time than a 3-0 result, which also makes sense due to time wasting in goal celebrations.

Regarding effective playing time, VAR impact is shown to be smaller, but still positive and significant, with the same control variables impacting the metric in the same direction.

Event studies, GLM and RDD confirm the increase in playing time due to VAR introduction. Regarding effective playing time, a bandwidth of 2 years when running an RDD yield a significant reduction as a first impact of VAR Introduction. In fact, in 2018 the effective playing time reduced. A larger 4-year bandwidth in RDD shows no effect on VAR impact, which argues that more playing time does not assure more effective playing time, raising concerns about game management.

These findings for Portuguese Football are in line with the ones found by Han et al. (2020), Errekagorri et al. (2020), Gürler & Polat (2021) and Meneguite et al. (2022). Additionally, Carlos et al. (2019) and Lago-Peñas et al. (2020) found evidence for an increase in added time in Bundesliga and La Liga. However, none of these studies seem to have accounted for the difference between playing time and effective playing time. This difference influences the actual playing dynamics, rather than just total length of the match. Moreover, studying effective playing time might support future decisions on injury prevention and games schedule.

6.4. RQ 4- VAR Impact on Competitiveness

Finally, for RQ 4, measured by the Winner Margin, a first approach found a slight increase in it, which would decrease the competitiveness of a league.

Advanced statistical methods concluded that this slight difference is hardly significant, giving more credits to home wins and total goals to decrease competitiveness, which indicates the natural expectation that home larger wins have little to no added competitiveness. It is worth noting that Playing Time did not add explanatory power to these research question. In fact, it worsens the model and some of the metrics, even changing intercept to negative, which makes no practical sense considering the Winner Margin definition.

GLM and RDD reveal a non-significant increase in winner margin following the introduction of the VAR. This finding can reassure football fans concerns about VAR potentially diminishing the competitiveness of matches. This implies that the fundamental competitive nature of football remains largely unaffected by the implementation of VAR technology.

Competitiveness in football has been sparsely studied, typically gauged by the diversity of league champions over the years. With no direct comparative studies, this research stands out in its unique approach to assessing competitiveness, paving the way for future studies and the exploration of new metrics in the field.

7. Conclusion

7.1. Developed work

This research delved into the impact of the VAR system on the dynamics of Portuguese football's first league. It encompassed a comprehensive examination of various aspects: goal scoring, home advantage, playing time, and overall competitiveness. The study's methodology evolved from descriptive analyses and non-parametric tests to event studies employing 13 distinct models. These included control variables, year fixed effects, and interaction terms, all further validated with Generalized Linear Models and Regression Discontinuity Design analyses.

Key findings from the study are as follows:

- VAR implementation marginally increased the average goals per game, with more emphasis to away teams.
- Home advantage remained largely unaffected by VAR.
- Playing time showed a significant increase post-VAR introduction and so did effective playing time, although at a considerably lower extent.
- VAR did not significantly influence the competitive balance of the game, as seen in the unchanged winner margin.

This research is pioneering in several aspects. It offers a thorough and robust examination of home advantage (HA) through multiple metrics. It innovates by focusing on effective playing time, rather than just total playing time, and explores VAR's impact on the game competitiveness.

Having evidence of positive correlation between playing time and players injuries, this findings for Portuguese League should alert team departments, especially considering the increasing frequency of matches in a season, and its potential implications for player career longevity.

Additionally, the observation that effective playing time isn't keeping pace with the overall increase in playing time raises questions about the efficacy of bigger extra times. Referees should be aware that simply adding more time at the end of games may be experiencing diminishing returns, a factor that warrants further attention.

7.2. Limitations and future research

This study focuses on Portugal's first division, which also limits its comparability with broader research, suggesting future studies should expand to include multiple leagues and divisions. Incorporating a diverse range of leagues would allow for more sophisticated analytical techniques, such as two-way fixed effects models in a difference-in-differences framework. Besides, going deeper on robustness methods GLM and RDD, as exploring it with year fixed effects and interactions terms would improve confidence in results.

Another aspect to consider is the potential masking effect of the Covid-19 pandemic on VAR's impact, given that the pandemic's influence on football dynamics has been well-documented. However, there were two seasons with VAR prior to the pandemic which still allowed robust results.

Expanding the time frame of analysis beyond the 2020/2021 season and including referee experience as a control variable, for instance, could also enhance future studies. Moreover, differentiating between matches with active VAR interventions and analysing the specific teams involved could provide a clearer picture of VAR's true impact.

Exploring other metrics to gauge competitiveness, such as the rate of underdog victories or the frequency of comebacks, can add depth to the understanding of match dynamics and unpredictability.

Finally, a more detailed examination of the contradictory impacts of winner margin and total goals on home advantage could be insightful. This paradox could be rooted in the observation that unbalanced games might favour the home team, whereas high-scoring matches could benefit away teams.

References

Albanese, A., Baert, S., & Verstraeten, O. (2020). Twelve eyes see more than eight. Referee bias and the introduction of additional assistant referees in soccer. PLOS ONE, 15(2), e0227758. https://doi.org/10.1371/journal.pone.0227758

Altmann, S., Forcher, L., Woll, A., & Härtel, S. (2023). Effective playing time affects physical match performance in soccer: An analysis according to playing position. Biology of Sport. https://doi.org/10.5114/biolsport.2023.123320

Anders, A., & Rotthoff, K. W. (2014). Is home-field advantage driven by the fans? Evidence from across the ocean. Applied Economics Letters, 21(16), 1165–1168. https://doi.org/10.1080/13504851.2014.914139

Armatas, V., & Pollard, R. (2012). Home advantage in Greek football. European Journal of Sport Science, 14(2), 116–122. https://doi.org/10.1080/17461391.2012.736537

Balmer, N. J., Nevill, A. M., & Williams, A. M. (2001). Home advantage in the Winter Olympics (1908-1998). Journal of Sports Sciences, 19(2), 129–139. https://doi.org/10.1080/026404101300036334

Balmer, N. J., Nevill, A. M., Lane, A. M., Ward, P., Williams, A. M., & Fairclough, S. (2007). Influence of crowd noise on soccer refereeing consistency in soccer. Wlv.openrepository.com. http://hdl.handle.net/2436/35332

Barnett, V., & Hilditch, S. (1993). The Effect of an Artificial Pitch Surface on Home Team Performance in Football (Soccer). Journal of the Royal Statistical Society. Series a (Statistics in Society), 156(1), 39. https://doi.org/10.2307/2982859

Boheim, R., Grübl, D., & Lackner, M. (2018). Choking Under Pressure: Evidence of the Causal Effect of Audience Size on Performance. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3249875

Bouke Klein Teeselink, van, & Dennie van Dolder. (2023). Does Losing Lead to Winning? An Empirical Analysis for Four Sports. Management Science, 69(1), 513–532. https://doi.org/10.1287/mnsc.2022.4372

Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. Journal of Sports Sciences, 25(11), 1185–1194. https://doi.org/10.1080/02640410601038576

Bryson, A., Dolton, P., Reade, J. J., Schreyer, D., & Singleton, C. (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. Economics Letters, 198, 109664. https://doi.org/10.1016/j.econlet.2020.109664

Buraimo, B., Simmons, R., & Maciaszczyk, M. (2011). Favoritism and referee bias in european soccer: evidence from the spanish league and the uefa champions league. Contemporary economic policy, 30(3), 329–343. Https://doi.org/10.1111/j.1465-7287.2011.00295.x

Carlos, L.-P., Ezequiel, R., & Anton, K. (2019). How does Video Assistant Referee (VAR) modify the game in elite soccer? International Journal of Performance Analysis in Sport, 19(4), 646–653. https://doi.org/10.1080/24748668.2019.1646521

Castellano, J., Blanco-Villaseñor, A., & Álvarez, D. (2011). Contextual Variables and Time-Motion Analysis in Soccer. International Journal of Sports Medicine, 32(06), 415–421. https://doi.org/10.1055/s-0031-1271771

CIES Football Observatory. (2019, October 28). Weekly Post: Effective playing time: records per club and league. CIES Football Observatory. https://football-observatory.com/Effective-playing-time-records-per-club-and

CIES Football Observatory. (2020, September 14). Weekly Post: What about the home advantage after the COVID-19 pandemic?. CIES Football Observatory. https://football-observatory.com/IMG/sites/b5wp/2020/wp304/en/

CIES Football Observatory. (2021, April). Pan European analysis on the fluidity of football matches. CIES Football Observatory. https://football-observatory.com/IMG/sites/mr/mr64/en/

Constantinou, A. C., Fenton, N. E., & Hunter Pollock, L. J. (2014). Bayesian networks for unbiased assessment of referee bias in Association Football. Psychology of Sport and Exercise, 15(5), 538–547. https://doi.org/10.1016/j.psychsport.2014.05.009

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. Journal of Sport & Exercise Psychology, 14(1), 13-27. https://doi.org/10.1123/jsep.14.1.13

Cueva, C. (2020). Animal Spirits in the Beautiful Game. Testing social pressure in professional football during the COVID-19 lockdown. https://doi.org/10.31219/osf.io/hczkj

Cunningham, S. (2021). CAUSAL INFERENCE: the mixtape. Yale University Press.

Dilger, A., & Vischer, L. (2020). No Home Bias in Ghost Games. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3674876

Dohmen, T. J. (2008a). Do professionals choke under pressure? Journal of Economic Behavior & Organization, 65(3-4), 636–653. https://ideas.repec.org/a/eee/jeborg/v65y2008i3-4p636-653.html

Dohmen, T. J. (2008b). THE INFLUENCE OF SOCIAL FORCES: EVIDENCE FROM THE BEHAVIOR OF FOOTBALL REFEREES. Economic Inquiry, 46(3), 411–424. https://doi.org/10.1111/j.1465-7295.2007.00112.x

Dvorak, J., Junge, A., Graf-Baumann, T., & Peterson, L. (2004). Football is the most popular sport worldwide. The American Journal of Sports Medicine, 32(1), 3–4. https://doi.org/10.1177/0363546503262283

Endrich, M., & Gesche, T. (2020). Home-Bias in Referee Decisions: Evidence from “Ghost Matches” During the COVID-19 Pandemic. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3639500

Erikstad, M. K., & Johansen, B. T. (2020). Referee Bias in Professional Football: Favoritism Toward Successful Teams in Potential Penalty Situations. Frontiers in Sports and Active Living, 2. https://doi.org/10.3389/fspor.2020.00019

Errekagorri, I., Castellano, J., Echeazarra, I., & Lago-Peñas, C. (2020). The effects of the Video Assistant Referee system (VAR) on the playing time, technical-tactical and physical performance in elite soccer. International Journal of Performance Analysis in Sport, 20(5), 808–817. https://ideas.repec.org/a/taf/rpanxx/v20y2020i5p808-817.html

Ferraresi, M., & Gucciardi, G. (2020). Team performance and audience: experimental evidence from the football sector. Working Papers. https://ideas.repec.org/p/ipu/wpaper/94.html

Ferraresi, M., & Gucciardi, G. (2021). Who chokes on a penalty kick? Social environment and individual performance during Covid-19 times. Economics Letters, 203, 109868. https://doi.org/10.1016/j.econlet.2021.109868

FIFA. (n.d.). Organisation. Retrieved November 8, 2023, from https://www.fifa.com/about-fifa/organisation.

Fischer, K., & Haucap, J. (2021). Does Crowd Support Drive the Home Advantage in Professional Football? Evidence from German Ghost Games during the COVID-19 Pandemic. Journal of Sports Economics, 22(8), 152700252110265. https://doi.org/10.1177/15270025211026552

Fisher, B. (2023, July 28). Premier League referees to add on time lost for goal celebrations this season. The Guardian. https://www.theguardian.com/football/2023/jul/28/premier-league-referees-to-add-on-time-lost-for-goal-celebrations-this-season

Applied Regression Analysis and Generalized Linear Models. (2020, January 14). SAGE Publications Inc. https://us.sagepub.com/en-us/nam/applied-regression-analysis-and-generalized-linear-models/book237254

Football Results, Statistics & Soccer Betting Odds Data. (n.d.). Www.football-Data.co.uk. https://www.football-data.co.uk/data.php

Franchini, E., & Takito, M. Y. (2016). Home advantage in combat sports during the Olympic Games. Sport Sciences for Health, 12(3), 287–290. https://doi.org/10.1007/s11332-016-0286-9

Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2001, July 1). Favoritism Under Social Pressure. Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=276006

Goller, D., & Krumer, A. (2020). Let's meet as usual: Do games played on non-frequent days differ? Evidence from top European soccer leagues. European Journal of Operational Research, 286(2), 740–754. https://doi.org/10.1016/j.ejor.2020.03.062

Gómez, M. A., Pollard, R., & Luis-Pascual, J. C. (2011). Comparison of the home advantage in nine different professional team sports in Spain. Perceptual and Motor Skills, 113(1), 150-156. https://doi.org/10.2466/05.PMS.113.4.150-156

Goodall, S., Thomas, K., Harper, L. D., Hunter, R., Parker, P., Stevenson, E., West, D., Russell, M., & Howatson, G. (2017). The assessment of neuromuscular fatigue during 120 min of simulated soccer exercise. European Journal of Applied Physiology, 117(4), 687–697. https://doi.org/10.1007/s00421-017-3561-9

Goumas, C. (2014). Home advantage in Australian soccer. Journal of Science and Medicine in Sport, 17(1), 119–123. https://doi.org/10.1016/j.jsams.2013.02.014

Greve, H. R., Rudi, N., & Walvekar, A. (2019). Strategic rule breaking: Time wasting to win soccer games. PLOS ONE, 14(12), e0224150. https://doi.org/10.1371/journal.pone.0224150

Gürler, C., & Polat, V. (2021). Video Assistant Referee's Effect on Football: Turkish Super League Case. Revista Brasileira de Futsal E Futebol, 13(52), 118–124. https://doaj.org/article/13c111a8436e4432b9985ebfd98a70e6

Hägglund, M., Waldén, M., Magnusson, H., Kristenson, K., Bengtsson, H., & Ekstrand, J. (2013). Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA Champions League injury study. British Journal of Sports Medicine, 47(12), 738–742. https://doi.org/10.1136/bjsports-2013-092215

Hamsund, T., & Scelles, N. (2021). Fans' Perceptions towards Video Assistant Referee (VAR) in the English Premier League. Journal of Risk and Financial Management, 14(12), 573. https://doi.org/10.3390/jrfm14120573

Han, B., Chen, Q., Lago-Peñas, C., Wang, C., & Liu, T. (2020). The influence of the video assistant referee on the Chinese Super League. International Journal of Sports Science & Coaching, 15(5-6), 662–668. https://doi.org/10.1177/1747954120938984

Harb-Wu, K., & Krumer, A. (2019). Choking under pressure in front of a supportive audience: Evidence from professional biathlon. Journal of Economic Behavior & Organization, 166, 246–262. https://doi.org/10.1016/j.jebo.2019.09.001

Harper, L. D., West, D. J., Stevenson, E., & Russell, M. (2014). Technical Performance Reduces during the Extra-Time Period of Professional Soccer Match-Play. PLoS ONE, 9(10), e110995. https://doi.org/10.1371/journal.pone.0110995

Holder, U., Ehrmann, T., & König, A. (2021). Monitoring experts: insights from the introduction of video assistant referee (VAR) in elite football. Journal of Business Economics, 92. https://doi.org/10.1007/s11573-021-01058-5

Hon, L. Y., & Parinduri, R. A. (2014). Does the Three-Point Rule Make Soccer More Exciting? Evidence From a Regression Discontinuity Design. Journal of Sports Economics, 17(4), 377–395. https://doi.org/10.1177/1527002514531790

Huntington-Klein, N. (2021). The Effect. CRC Press.

IFAB. (1970). Minutes of the Annual General Meeting. Retrieved from https://web.archive.org/web/20110430093219/http://ssbra.org/html/laws/IFABarc/pdf/1970/1970min.pdf

IFAB. (n.d.). Laws of the Game. Retrieved November 8, 2023, from https://www.theifab.com/laws

Inan, T. (2018). Analyzing the Home-Field Advantage in Major European Football Leagues. International Journal of Environmental and Science Education. http://www.ijese.net/makale/2011.html

Inan, T. (2020). The Effect of Crowd Support on Home-Field Advantage: Evidence from European Football. Annals of Applied Sport Science, 8(3). https://doi.org/10.29252/aassjournal.806

Jamieson, J. P. (2010). The Home Field Advantage in Athletics: A Meta-Analysis. Journal of Applied Social Psychology, 40(7), 1819–1848. https://doi.org/10.1111/j.1559-1816.2010.00641.x

Jones, M. B. (2013). The home advantage in individual sports: An augmented review. Psychology of Sport and Exercise, 14(3), 397–404. https://doi.org/10.1016/j.psychsport.2013.01.002

Kim, C. H., Lee, K. Y., & Kwon, Y. S. (2022). Does the Video Assistant Referee (VAR) technology mitigate referee bias in professional football leagues? Www.econstor.eu; Calgary: International Telecommunications Society (ITS). http://hdl.handle.net/10419/265643

Kolbinger, O., & Knopp, M. (2020). Video kills the sentiment—Exploring fans' reception of the video assistant referee in the English premier league using Twitter data. PLOS ONE, 15(12), e0242728. https://doi.org/10.1371/journal.pone.0242728

Kolbinger, O., & Link, D. (2016). The use of vanishing spray reduces the extent of rule violations in soccer. SpringerPlus, 5(1). https://doi.org/10.1186/s40064-016-3274-2

Kubayi, A., Larkin, P., & Toriola, A. (2021). The impact of video assistant referee (VAR) on match performance variables at men's FIFA World Cup tournaments. Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 236(3), 175433712199758. https://doi.org/10.1177/1754337121997581

Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. Korean Journal of Anesthesiology, 70(2), 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Lago-Peñas, C., & Gómez-López, M. (2016). The Influence of Referee Bias on Extra Time in Elite Soccer Matches. Perceptual and Motor Skills, 122(2), 666–677. https://doi.org/10.1177/0031512516633342

Lago-Peñas, C., Gómez, M., & Pollard, R. (2020). The effect of the Video Assistant Referee on referee's decisions in the Spanish LaLiga. International Journal of Sports Science & Coaching, 174795412098011. https://doi.org/10.1177/1747954120980111

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. Journal of Economic Literature, 48(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Leite, W. S. S. (2017). Home Advantage: Comparison between the Major European Football Leagues. Athens Journal of Sports, 4(1), 65–74. https://doi.org/10.30958/ajspo.4.1.4

Leite, W., & Pollard, R. (2018). International comparison of differences in home advantage between level 1 and level 2 of domestic football leagues. German Journal of Exercise and Sport Research, 48(2), 271–277. https://doi.org/10.1007/s12662-018-0507-2

Leitner, M. C., & Richlan, F. (2021). No Fans–No Pressure: Referees in Professional Football During the COVID-19 Pandemic. Frontiers in Sports and Active Living, 3. https://doi.org/10.3389/fspor.2021.720488

Leitner, M. C., Daumann, F., Follert, F., & Richlan, F. (2022). The cauldron has cooled down: a systematic literature review on home advantage in football during the COVID-19 pandemic from a socio-economic and psychological perspective. Management Review Quarterly. https://doi.org/10.1007/s11301-021-00254-5

Lesik, S. A. (2008). Studying the Effectiveness of Programs and Initiatives in Higher Education Using the Regression-Discontinuity Design. Springer EBooks, 277–297. https://doi.org/10.1007/978-1-4020-6959-8_9

Lidor, R., Bar-Eli, M., Arnon, M., & Bar-Eli, A. A. (2010). On the advantage of playing the second game at home in the knock out stages of European soccer cup competitions. International Journal of Sport and Exercise Psychology, 8(3), 312–325. https://doi.org/10.1080/1612197x.2010.9671956

Linke, D., Link, D., Weber, H., & Lames, M. (2018). Decline in Match Running Performance in Football is affected by an Increase in Game Interruptions. Journal of Sports Science & Medicine, 17(4), 662–667. https://pubmed.ncbi.nlm.nih.gov/30479536/

Magee, C., & Wolaver, A. (2023). Crowds and the Timing of Goals and Referee Decisions. Journal of Sports Economics, 152700252311607. https://doi.org/10.1177/15270025231160761

Marek, P., & Vávra, F. (2020). Comparison of Home Advantage in European Football Leagues. Risks, 8(3), 87. https://doi.org/10.3390/risks8030087

Matos, R., Amaro, N., & Pollard, R. (n.d.). How best to quantify home advantage in team sports: an investigation involving male senior handball leagues in Portugal and Spain. https://repositorio.ipsantarem.pt/bitstream/10400.15/3736/1/05902.pdf

McCarrick, D., Bilalic, M., Neave, N., & Wolfson, S. (2021). Home advantage during the COVID-19 pandemic: Analyses of European football leagues. Psychology of Sport and Exercise, 56(1), 102013. https://doi.org/10.1016/j.psychsport.2021.102013

Meneguite, Y. N., Leite, L. B., Silva, D. C., Moura, A. G., & Lavorato, V. N. (2022). Influence of the video assistant referee (VAR) on the Brazilian Men's Soccer Championship. Journal of Physical Education and Sport, 22(4), 858-862. https://doi.org/10.7752/jpes.2022.04108

Moffat, J. (2019). The impact of participation in pan-European competition on domestic performance in association football. European Sport Management Quarterly, 1–18. https://doi.org/10.1080/16184742.2019.1629980

Morita, H., & Araki, S. (2022). Social pressure in football matches: an event study of “Remote Matches” in Japan. Applied Economics Letters, 1–4. https://doi.org/10.1080/13504851.2022.2066617

Nevill, A. M., & Holder, R. L. (1999). Home Advantage in Sport. Sports Medicine, 28(4), 221–236. https://doi.org/10.2165/00007256-199928040-00001

Odds Portal. (n.d.).: Odds Comparison, Sports Betting Odds. Oddsportal.com. https://www.oddsportal.com/

Page, K., & Page, L. (2010). Alone against the crowd: Individual differences in referees’ ability to cope under pressure. Journal of Economic Psychology, 31(2), 192–199. https://doi.org/10.1016/j.joep.2009.08.007

Page, L., & Page, K. (2007). The second leg home advantage: Evidence from European football cup competitions. Journal of Sports Sciences, 25(14), 1547–1556. https://doi.org/10.1080/02640410701275219

Peeters, T., & van Ours, J. C. (2020). Seasonal Home Advantage in English Professional Football; 1974–2018. De Economist. https://doi.org/10.1007/s10645-020-09372-z

Pettersson-Lidbom, P., & Priks, M. (2010). Behavior under social pressure: Empty Italian stadiums and referee bias. Economics Letters, 108(2), 212–214. https://doi.org/10.1016/j.econlet.2010.04.023

Picazo-Tadeo, A. J., González-Gómez, F., & Guardiola, J. (2016). Does the crowd matter in refereeing decisions? Evidence from Spanish soccer. International Journal of Sport and Exercise Psychology, 15(5), 447–459. https://doi.org/10.1080/1612197x.2015.1126852

Plumley, D., Ramchandani, G., & Wilson, R. (2018). Mind the gap: an analysis of competitive balance in the English Football League system. International Journal of Sport Management and Marketing, 18(5), 357. https://doi.org/10.1504/ijsmm.2018.094344

Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. Journal of Sports Sciences, 4(3), 237–248. https://doi.org/10.1080/02640418608732122

Pollard, R. (2002). Evidence of a reduced home advantage when a team moves to a new stadium. Journal of Sports Sciences, 20(12), 969–973. https://doi.org/10.1080/026404102321011724

Pollard, R. (2006). Worldwide regional variations in home advantage in association football. Journal of Sports Sciences, 24(3), 231–240. https://doi.org/10.1080/02640410500141836

Pollard, R. (2008). Home Advantage in Football: A Current Review of an Unsolved Puzzle. The Open Sports Sciences Journal, 1(1), 12–14. https://doi.org/10.2174/1875399x00801010012

Pollard, R., & Gómez, M. A. (2012). Comparison of home advantage in men's and women's football leagues in Europe. European Journal of Sport Science, 14(sup1), S77–S83. https://doi.org/10.1080/17461391.2011.651490

Pollard, R., & Gómez, M. A. (2014). Components of home advantage in 157 national soccer leagues worldwide. International Journal of Sport and Exercise Psychology, 12(3), 218–233. https://doi.org/10.1080/1612197x.2014.888245

Pollard, R., & Gómez, M. A. (2015). Comparison of Home Advantage in College and Professional Team Sports in the United States. Collegium Antropologicum, 39(3), 583–589. https://pubmed.ncbi.nlm.nih.gov/26898053/

Pollard, R., Prieto, J., & Gómez, M.-Á. (2017). Global differences in home advantage by country, sport and sex. International Journal of Performance Analysis in Sport, 17(4), 586–599. https://doi.org/10.1080/24748668.2017.1372164

Pollard, R., Prieto, J., & Gómez, M.-Á. (2017). Global differences in home advantage by country, sport and sex. International Journal of Performance Analysis in Sport, 17(4), 586-599. https://doi.org/10.1080/24748668.2017.1372164

Ponzo, M., & Scoppa, V. (2016). Does the Home Advantage Depend on Crowd Support? Evidence From Same-Stadium Derbies. Journal of Sports Economics, 19(4), 562–582. https://doi.org/10.1177/1527002516665794

Psiuk, R., Seidl, T., Strauß, W., & Bernhard, J. (2014). Analysis of Goal Line Technology from the Perspective of an Electromagnetic Field based Approach. Procedia Engineering, 72, 279–284. https://doi.org/10.1016/j.proeng.2014.06.050

Ramchandani, G., Plumley, D., Boyes, S., & Wilson, R. (2018). A longitudinal and comparative analysis of competitive balance in five European football leagues. Team Performance Management: An International Journal, 24(5/6), 265–282. https://doi.org/10.1108/tpm-09-2017-0055

Ramchandani, G., Plumley, D., Mondal, S., Millar, R., & Wilson, R. (2023). “You can look, but don’t touch”: competitive balance and dominance in the UEFA Champions League. Soccer & Society, 1–13. https://doi.org/10.1080/14660970.2023.2194512

Rampinini, E., Impellizzeri, F. M., Castagna, C., Coutts, A. J., & Wisløff, U. (2009). Technical performance during soccer matches of the Italian Serie A league: effect of fatigue and competitive level. Journal of Science and Medicine in Sport, 12(1), 227–233. https://doi.org/10.1016/j.jsams.2007.10.002

Reade, J. J., Schreyer, D., & Singleton, C. (2021). Eliminating supportive crowds reduces referee bias. Economic Inquiry. https://doi.org/10.1111/ecin.13063

Reilly, B., & Witt, R. (2021). The Effect of League Design on Spectator Attendance: A Regression Discontinuity Design Approach. Journal of Sports Economics, 152700252198939. https://doi.org/10.1177/1527002521989393

Rey, E., Kalén, A., Lorenzo-Martínez, M., López-Del Campo, R., Nevado-Garrosa, F., & Lago-Peñas, C. (2020). Elite Soccer Players do Not Cover Less Distance in the Second Half of the Matches When Game Interruptions Are Considered. Journal of Strength and Conditioning Research, Publish Ahead of Print. https://doi.org/10.1519/jsc.0000000000003935

Ribeiro, H. V., Mukherjee, S., & Zeng, X. H. (2016). The advantage of playing home in NBA: Microscopic, team-specific and evolving features. PloS One, 11(3), e0152440. https://doi.org/10.1371/journal.pone.0152440

Riedl, D., Strauss, B., Heuer, A., & Rubner, O. (2014). Finale furioso: referee-biased injury times and their effects on home advantage in football. Journal of Sports Sciences, 33(4), 327–336. https://doi.org/10.1080/02640414.2014.944558

Russell, M., Sparkes, W., Northeast, J., & Kilduff, L. P. (2015). Responses to a 120 min reserve team soccer match: a case study focusing on the demands of extra time. Journal of Sports Sciences, 33(20), 2133–2139. https://doi.org/10.1080/02640414.2015.1064153

Sánchez, P. A., García-Calvo, T., Leo, F. M., Pollard, R., & Gómez, M. A. (2009). An Analysis of Home Advantage in the Top Two Spanish Professional Football Leagues. Perceptual and Motor Skills, 108(3), 789–797. https://doi.org/10.2466/pms.108.3.789-797

Sapp, R. M., Spangenburg, E. E., & Hagberg, J. M. (2017). Trends in aggressive play and refereeing among the top five European soccer leagues. Journal of Sports Sciences, 36(12), 1346–1354. https://doi.org/10.1080/02640414.2017.1377911

Scanlon, C., Griggs, G., & McGillick, C. (2022). “It’s not football anymore”: perceptions of the video assistant referee by english premier league football fans. Soccer & Society, 23(8), 1–13. https://doi.org/10.1080/14660970.2022.2033731

Scoppa, V. (2007). Are subjective evaluations biased by social factors or connections? An econometric analysis of soccer referee decisions. Empirical Economics, 35(1), 123–140. https://doi.org/10.1007/s00181-007-0146-1

Scoppa, V. (2020). Social Pressure in the Stadiums: Do Agents Change Behavior Without Crowd Support? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3674316

Soto-Valero, C., & Pic, M. (2019). Assessing the causal impact of the 3-point per victory scoring system in the competitive balance of LaLiga. International Journal of Computer Science in Sport, 18(3), 69–88. https://doi.org/10.2478/ijcss-2019-0018

Spitz, J., Wagemans, J., Memmert, D., Williams, A. M., & Helsen, W. F. (2020). Video assistant referees (VAR): The impact of technology on decision making in association football referees. Journal of Sports Sciences, 39(2), 1–7. https://doi.org/10.1080/02640414.2020.1809163

Staufenbiel, K., Lobinger, B., & Strauss, B. (2015). Home advantage in soccer – A matter of expectations, goal setting and tactical decisions of coaches? Journal of Sports Sciences, 33(18), 1932–1941. https://doi.org/10.1080/02640414.2015.1018929

Tilp, M., & Thaller, S. (2020). Covid-19 Has Turned Home Advantage Into Home Disadvantage in the German Soccer Bundesliga. Frontiers in Sports and Active Living, 2. https://doi.org/10.3389/fspor.2020.593499

UEFA. (2021, June 24). Abolition of the away goals rule in all UEFA club competitions. UEFA.com. https://it.uefa.com/returntoplay/news/026a-1298aeb73a7a-5b64cb68d920-1000--abolition-of-the-away-goals-rule-in-all-uefa-club-competi/

Vale, A. S. do. (2023, January 24). Analysis of Video Assistant Referee (VAR) in football. Run.unl.pt. http://hdl.handle.net/10362/148992

Van Damme, N., & Baert, S. (2019). Home advantage in European international soccer: Which dimension of distance matters? EconStor https://www.econstor.eu/handle/10419/191758

Van de Ven, N. (2011). Supporters Are Not Necessary for the Home Advantage: Evidence From Same-Stadium Derbies and Games Without an Audience. Journal of Applied Social Psychology, 41(12), 2785–2792. https://doi.org/10.1111/j.1559-1816.2011.00865.x

Van der Burg, T. (2023). Competitive balance and demand for European men's football: a review of the literature. Managing Sport and Leisure, 1–16. https://doi.org/10.1080/23750472.2023.2206815

Winand, M., Schneiders, C., Merten, S., & Marlier, M. (2021). Sports fans and innovation: An analysis of football fans' satisfaction with video assistant refereeing through social identity and argumentative theories. Journal of Business Research, 136(1), 99–109. https://doi.org/10.1016/j.jbusres.2021.07.029

Winder, N., Russell, M., Naughton, R., & Harper, L. (2018). The Impact of 120 Minutes of Match-Play on Recovery and Subsequent Match Performance: A Case Report in Professional Soccer Players. Sports, 6(1), 22. https://doi.org/10.3390/sports6010022

Wunderlich, F., Weigelt, M., Rein, R., & Memmert, D. (2021). How does spectator presence affect football? Home advantage remains in European top-class football matches played without spectators during the COVID-19 pandemic. PLOS ONE, 16(3), e0248590. https://doi.org/10.1371/journal.pone.0248590

Zhang, Y., Li, D., Gómez-Ruano, M.-Á., Memmert, D., Li, C., & Fu, M. (2022). The effect of the video assistant referee (VAR) on referees' decisions at FIFA Women's World Cups. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.984367

Appendix A - Descriptive Analysis

	TOTAL GOALS	HOME GOALS	AWAY GOALS	HOME ADVANTAGE	HOME WIN	HOME POINTS	WINNER MARGIN	ODD HOME	ODD AWAY	PLAYING TIME	EFFECTIVE PLAYING TIME
MEAN PRE VAR	2.4965	1.4240	1.0725	0.3515	0.4439	1.5915	1.3394	0.4313	0.3134	51.0131	0.5341
MEAN POST VAR	2.5773	1.4260	1.1513	0.2747	0.4522	1.5846	1.3573	0.4332	0.3130	53.7077	0.5416
MIN PRE VAR	0.0000	0.0000	0.0000	-6.0000	0.0000	0.0000	0.0000	0.0916	0.0235	28.4300	0.2942
MIN POST VAR	0.0000	0.0000	0.0000	-7.0000	0.0000	0.0000	0.0000	0.0530	0.0325	21.5200	0.2270
MAX PRE VAR	8.0000	7.0000	6.0000	7.0000	1.0000	3.0000	7.0000	0.9063	0.7960	66.1667	0.7018
MAX POST VAR	10.0000	10.0000	8.0000	10.0000	1.0000	3.0000	10.0000	0.8878	0.8329	80.5800	0.9082
25% PRE VAR	1.0000	0.0000	0.0000	-1.0000	0.0000	0.0000	0.0000	0.3173	0.1867	47.9500	0.5013
25% POST VAR	1.0000	1.0000	0.0000	-1.0000	0.0000	0.0000	1.0000	0.3172	0.1981	50.5850	0.5094
75% PRE VAR	4.0000	2.0000	2.0000	1.0000	1.0000	3.0000	2.0000	0.5418	0.3822	54.2775	0.5686
75% POST VAR	4.0000	2.0000	2.0000	1.0000	1.0000	3.0000	2.0000	0.5327	0.3786	56.8500	0.5754
SD PRE VAR	1.5644	1.2476	1.1062	1.7644	0.4971	1.3127	1.2006	0.1947	0.1799	4.7348	0.0515
SD POST VAR	1.6494	1.2532	1.1812	1.7919	0.4979	1.3372	1.2011	0.1959	0.1794	4.7162	0.0529
VAR PRE VAR	2.4473	1.5564	1.2238	3.1132	0.2471	1.7232	1.4413	0.0379	0.0324	22.4185	0.0027
VAR POST VAR	2.7205	1.5704	1.3953	3.2109	0.2479	1.7880	1.4426	0.0384	0.0322	22.2423	0.0028
95% CI LOW PRE VAR	2.4063	1.3521	1.0088	0.2497	0.4152	1.5158	1.2702	0.4201	0.3030	50.7401	0.5311
95% CI LOW POST VAR	2.4847	1.3557	1.0850	0.1742	0.4242	1.5096	1.2899	0.4223	0.3029	53.4431	0.5386
95% CI HIGH PRE VAR	2.5867	1.4959	1.1363	0.4532	0.4725	1.6672	1.4086	0.4425	0.3238	51.2860	0.5371
95% CI HIGH POST VAR	2.6698	1.4963	1.2175	0.3753	0.4801	1.6596	1.4247	0.4442	0.3231	53.9723	0.5445

Table A.1 Descriptive Statistics

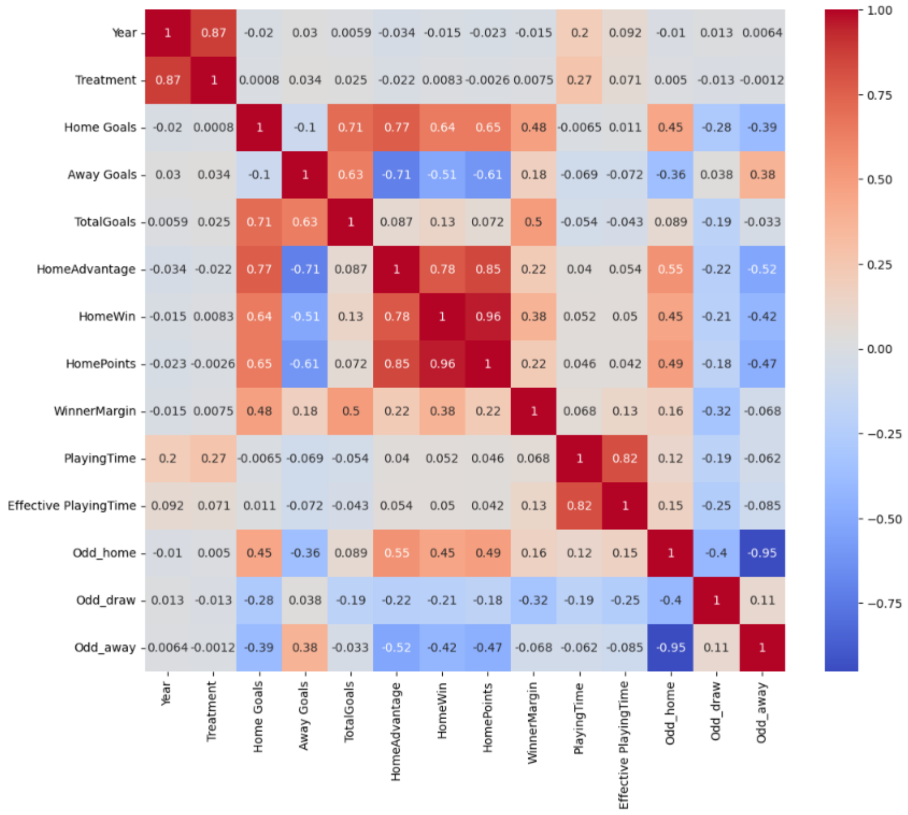

Figure 341. Raw Correlation Matrix

Appendix B- Event Studies

						Dependent variable: TotalGoals
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	2.4965***	2.1804***	2.3031***	3.5962***	1.5991***	3.2418***
	(0.0497)	(0.0904)	(0.0592)	(0.3612)	(0.0540)	(0.3044)
Treatment	0.0807	0.0793	0.0771	0.1388*	0.0687	0.1627***
	(0.0686)	(0.0685)	(0.0680)	(0.0710)	(0.0565)	(0.0587)
Odd_home		0.7331***				0.4452***
		(0.1725)				(0.1582)
HomeWin			0.4357***			-0.2761***
			(0.0652)			(0.0650)
PlayingTime				-0.0216***		-0.0346***
				(0.0069)		(0.0059)
WinnerMargin					0.6701***	0.7110***
					(0.0203)	(0.0196)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0006	0.0085	0.0188	0.0046	0.2507	0.2657
Adjusted R^2	0.0002	0.0077	0.0180	0.0038	0.2500	0.2641
Residual Std. Error	1.6086 (df=2379)	1.6026 (df=2378)	1.5943 (df=2378)	1.6057 (df=2378)	1.3932 (df=2378)	1.3800 (df=2375)
F Statistic	1.3842 (df=1; 2379)	9.6253*** (df=2; 2378)	23.0601*** (df=2; 2378)	5.5039*** (df=2; 2378)	543.2556*** (df=2; 2378)	273.6687*** (df=5; 2375)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.1. RQ 1 – Event Studies: Models 1-6

						Dependent variable: TotalGoals
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	0.1226	0.1192	0.1236	0.0842	0.0495	-0.0189
	(0.1173)	(0.1185)	(0.1160)	(0.1183)	(0.0906)	(0.0903)
C(Year)[T.2016]	0.3449^{**}	0.3410^{**}	0.3514^{**}	0.3161^{**}	0.2630^{**}	0.2054^{*}
	(0.1400)	(0.1411)	(0.1396)	(0.1390)	(0.1132)	(0.1112)
C(Year)[T.2017]	0.0083	0.0070	0.0106	-0.0126	0.0114	-0.0241
	(0.1235)	(0.1237)	(0.1215)	(0.1240)	(0.1004)	(0.1006)
C(Year)[T.2018]	0.3285^{***}	0.3196^{***}	0.3001^{***}	0.4023^{***}	0.1922^{**}	0.3143^{***}
	(0.1130)	(0.1136)	(0.1130)	(0.1156)	(0.0914)	(0.0918)
C(Year)[T.2019]	0.3308^{**}	0.3169^{**}	0.3241^{**}	0.3458^{**}	0.2262^{**}	0.2394^{**}
	(0.1439)	(0.1432)	(0.1433)	(0.1434)	(0.1154)	(0.1121)
C(Year)[T.2020]	0.1226	0.1191	0.1390	0.1463	0.1453	0.1717^{*}
	(0.1223)	(0.1229)	(0.1208)	(0.1220)	(0.0947)	(0.0939)
C(Year)[T.2021]	0.0442	0.0560	0.0592	0.0777	0.0538	0.1054
	(0.1089)	(0.1099)	(0.1080)	(0.1087)	(0.0876)	(0.0867)
Intercept	2.3708^{***}	2.0666^{***}	2.1778^{***}	3.5258^{***}	1.5182^{***}	3.2435^{***}
	(0.0817)	(0.1115)	(0.0879)	(0.3741)	(0.0668)	(0.3144)
Odd_home		0.7106^{***}				0.4382^{***}
		(0.1730)				(0.1581)
HomeWin			0.4289^{***}			-0.2769^{***}
			(0.0655)			(0.0652)
PlayingTime				-0.0222^{***}		-0.0354^{***}
				(0.0070)		(0.0059)
WinnerMargin					0.6666^{***}	0.7077^{***}
					(0.0204)	(0.0198)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0075	0.0149	0.0250	0.0116	0.2537	0.2687
Adjusted R^2	0.0046	0.0116	0.0217	0.0083	0.2512	0.2653
Residual Std. Error	1.6051 (df=2373)	1.5994 (df=2372)	1.5912 (df=2372)	1.6021 (df=2372)	1.3922 (df=2372)	1.3790 (df=2369)
F Statistic	2.6445^{**} (df=7; 2373)	4.3399^{***} (df=8; 2372)	7.6430^{***} (df=8; 2372)	3.3144^{***} (df=8; 2372)	140.9493^{***} (df=8; 2372)	129.0114^{***} (df=11; 2369)
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B.2. RQ 1 - Event Studies: Models 7-12

	Dependent variable: TotalGoals
	Model 13
HomeWin	-0.2446***
	(0.0925)
Intercept	3.2398***
	(0.4239)
Odd_home	0.4880**
	(0.2110)
PlayingTime	-0.0347***
	(0.0083)
Treatment	0.1736
	(0.6261)
Treatment:HomeWin	-0.0556
	(0.1302)
Treatment:Odd_home	-0.0791
	(0.3168)
Treatment:PlayingTime	-0.0001
	(0.0118)
Treatment:WinnerMargin	0.0399
	(0.0390)
WinnerMargin	0.6894***
	(0.0274)
N	2381.0
Observations	2381
R^2	0.2659
Adjusted R^2	0.2631
Residual Std. Error	1.3810 (df=2371)
F Statistic	156.6965*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.3. RQ 1 - Event Studies: Model 13

						Dependent variable: HomeAdvantage
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	0.3515***	1.9718***	0.1086	-0.5604	-0.0795	1.6311***
	(0.0447)	(0.0735)	(0.0739)	(0.3824)	(0.0699)	(0.3331)
Treatment	-0.0767	-0.0789	-0.0846	-0.1249*	-0.0825	-0.0810
	(0.0684)	(0.0658)	(0.0678)	(0.0709)	(0.0667)	(0.0637)
Odd_away		-5.1699***				-5.0478***
		(0.1929)				(0.2125)
TotalGoals			0.0973***			-0.0301
			(0.0300)			(0.0190)
PlayingTime				0.0179**		-0.0002
				(0.0075)		(0.0059)
WinnerMargin					0.3218***	0.2910***
					(0.0570)	(0.0441)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0005	0.2731	0.0082	0.0027	0.0476	0.3069
Adjusted R^2	0.0000	0.2725	0.0074	0.0019	0.0468	0.3054
Residual Std. Error	1.7786 (df=2379)	1.5171 (df=2378)	1.7721 (df=2378)	1.7770 (df=2378)	1.7365 (df=2378)	1.4823 (df=2375)
F Statistic	1.2583 (df=1; 2379)	359.0632*** (df=2; 2378)	6.6343*** (df=2; 2378)	3.5354** (df=2; 2378)	18.3518*** (df=2; 2378)	160.8640*** (df=5; 2375)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.4. RQ 2.1 - Event Studies: Models 1-6

	Dependent variable: HomeAdvantage					
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	0.1210	0.0919	0.1093	0.1509	0.0860	0.0621
	(0.1245)	(0.1213)	(0.1242)	(0.1248)	(0.1213)	(0.1165)
C(Year)[T.2016]	0.0230	-0.0002	-0.0099	0.0454	-0.0162	-0.0265
	(0.1341)	(0.1339)	(0.1342)	(0.1350)	(0.1315)	(0.1312)
C(Year)[T.2017]	0.0981	0.0598	0.0973	0.1144	0.0996	0.0610
	(0.1257)	(0.1157)	(0.1254)	(0.1268)	(0.1225)	(0.1126)
C(Year)[T.2018]	0.1504	0.1056	0.1191	0.0929	0.0852	0.0621
	(0.1373)	(0.1328)	(0.1375)	(0.1389)	(0.1358)	(0.1305)
C(Year)[T.2019]	0.0535	-0.0296	0.0219	0.0418	0.0034	-0.0622
	(0.1392)	(0.1284)	(0.1383)	(0.1397)	(0.1338)	(0.1228)
C(Year)[T.2020]	-0.0882	-0.1114	-0.0998	-0.1066	-0.0773	-0.0958
	(0.1437)	(0.1405)	(0.1436)	(0.1435)	(0.1441)	(0.1385)
C(Year)[T.2021]	-0.1666	-0.1199	-0.1708	-0.1927	-0.1620	-0.1134
	(0.1285)	(0.1272)	(0.1282)	(0.1292)	(0.1275)	(0.1238)
Intercept	0.2875^{***}	1.9291^{***}	0.0615	-0.6126	-0.1205	1.6667^{***}
	(0.0936)	(0.1099)	(0.1107)	(0.4105)	(0.1094)	(0.3595)
Odd_away		-5.1616^{***}				-5.0448^{***}
		(0.1929)				(0.2125)
PlayingTime				0.0173^{**}		-0.0014
				(0.0076)		(0.0060)
TotalGoals			0.0953^{***}			-0.0305
			(0.0301)			(0.0192)
WinnerMargin					0.3190^{***}	0.2896^{***}
					(0.0573)	(0.0444)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0033	0.2746	0.0107	0.0053	0.0494	0.3079
Adjusted R^2	0.0004	0.2722	0.0073	0.0020	0.0462	0.3046
Residual Std. Error	1.7783 (df=2373)	1.5174 (df=2372)	1.7721 (df=2372)	1.7769 (df=2372)	1.7370 (df=2372)	1.4832 (df=2369)
F Statistic	1.3780 (df=7; 2373)	92.0483^{***} (df=8; 2372)	2.6494^{***} (df=8; 2372)	1.8593^{*} (df=8; 2372)	5.8771^{***} (df=8; 2372)	75.8228^{***} (df=11; 2369)
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B.5. RQ 2.1 - Event Studies: Models 7-12

	Dependent variable: HomeAdvantage
	Model 13
Intercept	1.2560**
	(0.4899)
Odd_away	-4.9864***
	(0.3211)
PlayingTime	0.0053
	(0.0089)
TotalGoals	-0.0265
	(0.0275)
Treatment	0.6404
	(0.6651)
Treatment:Odd_away	-0.0968
	(0.4268)
Treatment:PlayingTime	-0.0106
	(0.0118)
Treatment:TotalGoals	-0.0056
	(0.0379)
Treatment:WinnerMargin	-0.0909
	(0.0875)
WinnerMargin	0.3375***
	(0.0537)
N	2381.0
Observations	2381
R^2	0.3081
Adjusted R^2	0.3055
Residual Std. Error	1.4822 (df=2371)
F Statistic	92.2829*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.6. RQ 2.1 - Event Studies: Model 13

	Dependent variable: HomeWin					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-0.2255***	1.6091***	-0.6541***	-1.3491***	-1.2517***	1.3323**
	(0.0526)	(0.1104)	(0.0832)	(0.4365)	(0.0839)	(0.5683)
Treatment	0.0336	0.0438	0.0205	-0.0257	0.0331	0.0606
	(0.0791)	(0.0936)	(0.0799)	(0.0804)	(0.0856)	(0.1012)
Odd_away		-6.2599***				-7.0100***
		(0.3687)				(0.4200)
PlayingTime				0.0220***		-0.0093
				(0.0084)		(0.0102)
TotalGoals			0.1704***			-0.1211***
			(0.0262)			(0.0342)
WinnerMargin					0.7695***	1.0060***
					(0.0551)	(0.0577)
N	2381	2381	2381	2381	2381	2381
Observations	2381	2381	2381	2381	2381	2381
Note:				*p<0.1; **p<0.05; ***p<0.01		

Table B.7. RQ 2.2 - Event Studies: Models 1-6

						Dependent variable: HomeWin
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	-0.0092	-0.0617	-0.0304	0.0248	-0.0764	-0.1182
	(0.1473)	(0.1843)	(0.1482)	(0.1473)	(0.1622)	(0.1971)
C(Year)[T.2016]	-0.0623	-0.1273	-0.1226	-0.0370	-0.1441	-0.2022
	(0.1508)	(0.1866)	(0.1544)	(0.1515)	(0.1672)	(0.2053)
C(Year)[T.2017]	-0.0225	-0.0789	-0.0241	-0.0039	-0.0153	-0.0668
	(0.1509)	(0.1790)	(0.1508)	(0.1516)	(0.1591)	(0.1870)
C(Year)[T.2018]	0.2661	0.2443	0.2153	0.2011	0.1799	0.2203
	(0.1657)	(0.2043)	(0.1690)	(0.1673)	(0.1805)	(0.2221)
C(Year)[T.2019]	0.0627	-0.0280	0.0081	0.0495	-0.0221	-0.0645
	(0.1554)	(0.1820)	(0.1588)	(0.1556)	(0.1688)	(0.1962)
C(Year)[T.2020]	-0.1560	-0.2261	-0.1797	-0.1773	-0.1261	-0.1663
	(0.1484)	(0.1808)	(0.1493)	(0.1489)	(0.1616)	(0.2021)
C(Year)[T.2021]	-0.1425	-0.1065	-0.1523	-0.1724	-0.1456	-0.1065
	(0.1560)	(0.1895)	(0.1571)	(0.1575)	(0.1685)	(0.2023)
Intercept	-0.2007*	1.6797***	-0.6035***	-1.2262***	-1.1874***	1.6426***
	(0.1077)	(0.1644)	(0.1215)	(0.4586)	(0.1297)	(0.5858)
Odd_away		-6.2579***				-7.0217***
		(0.3683)				(0.4210)
PlayingTime				0.0197**		-0.0133
				(0.0086)		(0.0103)
TotalGoals			0.1689***			-0.1223***
			(0.0265)			(0.0345)
WinnerMargin					0.7672***	1.0068***
					(0.0554)	(0.0580)
N	2381	2381	2381	2381	2381	2381
Observations	2381	2381	2381	2381	2381	2381
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.8. RQ 2.2 - Event Studies: Models 7-12

	Dependent variable: HomeWin
	Model 13
Intercept	0.6691***
	(0.1306)
Odd_away	-1.0444***
	(0.0652)
PlayingTime	-0.0015
	(0.0024)
TotalGoals	-0.0214**
	(0.0086)
Treatment	-0.0245
	(0.1887)
Treatment:Odd_away	-0.1054
	(0.0863)
Treatment:PlayingTime	0.0020
	(0.0034)
Treatment:TotalGoals	-0.0018
	(0.0118)
Treatment:WinnerMargin	-0.0255*
	(0.0148)
WinnerMargin	0.1741***
	(0.0100)
N	2381.0
Observations	2381
R^2	0.3056
Adjusted R^2	0.3030
Residual Std. Error	0.4153 (df=2371)
F Statistic	230.3707*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.9. RQ 2.2 - Event Studies - Model 13

						Dependent variable: HomePoints
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	1.5915***	0.1709***	1.4443***	0.8926***	1.2667***	0.4133*
	(0.0342)	(0.0565)	(0.0509)	(0.2866)	(0.0395)	(0.2489)
Treatment	-0.0069	-0.0133	-0.0117	-0.0438	-0.0113	0.0044
	(0.0520)	(0.0491)	(0.0519)	(0.0529)	(0.0508)	(0.0479)
Odd_home		3.2938***				3.1550***
		(0.0982)				(0.1075)
PlayingTime				0.0137**		-0.0063
				(0.0055)		(0.0049)
TotalGoals			0.0590***			-0.0487***
			(0.0161)			(0.0163)
WinnerMargin					0.2425***	0.1940***
					(0.0209)	(0.0186)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0000	0.2357	0.0051	0.0024	0.0483	0.2590
Adjusted R^2	-0.0004	0.2350	0.0043	0.0016	0.0475	0.2575
Residual Std. Error	1.3253 (df=2379)	1.1589 (df=2378)	1.3222 (df=2378)	1.3240 (df=2378)	1.2932 (df=2378)	1.1418 (df=2375)
F Statistic	0.0177 (df=1; 2379)	562.5926*** (df=2; 2378)	6.7280*** (df=2; 2378)	3.1132** (df=2; 2378)	67.8252*** (df=2; 2378)	366.8821*** (df=5; 2375)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.10. RQ 2.3 - Event Studies: Models 1-6

						Dependent variable: HomePoints
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	0.0209	0.0052	0.0138	0.0421	-0.0055	-0.0238
	(0.0947)	(0.0938)	(0.0952)	(0.0944)	(0.0943)	(0.0909)
C(Year)[T.2016]	-0.0477	-0.0657	-0.0676	-0.0318	-0.0772	-0.0827
	(0.1008)	(0.0992)	(0.1018)	(0.1010)	(0.0995)	(0.0973)
C(Year)[T.2017]	-0.0052	-0.0112	-0.0057	0.0063	-0.0041	-0.0175
	(0.0984)	(0.0930)	(0.0982)	(0.0986)	(0.0964)	(0.0906)
C(Year)[T.2018]	0.1484	0.1070	0.1294	0.1077	0.0992	0.1131
	(0.1102)	(0.1063)	(0.1112)	(0.1110)	(0.1091)	(0.1056)
C(Year)[T.2019]	-0.0000	-0.0643	-0.0191	-0.0083	-0.0377	-0.0702
	(0.1037)	(0.0972)	(0.1048)	(0.1035)	(0.1025)	(0.0953)
C(Year)[T.2020]	-0.1065	-0.1230	-0.1136	-0.1196	-0.0983	-0.1009
	(0.1004)	(0.0955)	(0.1004)	(0.1004)	(0.0999)	(0.0943)
C(Year)[T.2021]	-0.1033	-0.0488	-0.1058	-0.1217	-0.0998	-0.0334
	(0.1009)	(0.0979)	(0.1011)	(0.1014)	(0.1005)	(0.0967)
Intercept	1.6000^{***}	0.1910^{**}	1.4632^{***}	0.9630^{***}	1.2926^{***}	0.5506^{**}
	(0.0712)	(0.0838)	(0.0781)	(0.3007)	(0.0728)	(0.2589)
Odd_home		3.2911^{***}				3.1610^{***}
		(0.0985)				(0.1078)
PlayingTime				0.0122^{**}		-0.0083^{*}
				(0.0056)		(0.0049)
TotalGoals			0.0577^{***}			-0.0489^{***}
			(0.0163)			(0.0164)
WinnerMargin					0.2403^{***}	0.1929^{***}
					(0.0212)	(0.0186)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0033	0.2380	0.0082	0.0052	0.0505	0.2612
Adjusted R^2	0.0004	0.2354	0.0049	0.0018	0.0473	0.2578
Residual Std. Error	1.3248 (df=2373)	1.1586 (df=2372)	1.3218 (df=2372)	1.3239 (df=2372)	1.2933 (df=2372)	1.1416 (df=2369)
F Statistic	1.0872 (df=7; 2373)	146.7404^{***} (df=8; 2372)	2.7274^{***} (df=8; 2372)	1.5318 (df=8; 2372)	18.2933^{***} (df=8; 2372)	172.4132^{***} (df=11; 2369)
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B.11. RQ 2.3 - Event Studies: Models 7-12

	Dependent variable: HomePoints
	Model 13
Intercept	0.5446
	(0.3594)
Odd_home	3.0506***
	(0.1674)
PlayingTime	-0.0088
	(0.0071)
TotalGoals	-0.0444*
	(0.0247)
Treatment	-0.2795
	(0.5010)
Treatment:Odd_home	0.1959
	(0.2171)
Treatment:PlayingTime	0.0054
	(0.0098)
Treatment:TotalGoals	-0.0071
	(0.0329)
Treatment:WinnerMargin	-0.0489
	(0.0370)
WinnerMargin	0.2191***
	(0.0264)
N	2381.0
Observations	2381
R^2	0.2598
Adjusted R^2	0.2570
Residual Std. Error	1.1422 (df=2371)
F Statistic	213.5841*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.12. RQ 2.3 - Event Studies: Model 13

						Dependent variable: Odd_home
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	0.4313***	0.4101***	0.4043***	0.1797***	0.3961***	0.1947***
	(0.0039)	(0.0040)	(0.0073)	(0.0415)	(0.0063)	(0.0344)
Treatment	0.0019	0.0066	0.0011	-0.0113*	0.0015	-0.0045
	(0.0055)	(0.0056)	(0.0055)	(0.0061)	(0.0055)	(0.0057)
HomeAdvantage		0.0602***				0.0589***
		(0.0017)				(0.0016)
PlayingTime				0.0049***		0.0039***
				(0.0008)		(0.0007)
TotalGoals			0.0108***			0.0046*
			(0.0025)			(0.0024)
WinnerMargin					0.0263***	0.0032
					(0.0040)	(0.0030)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0000	0.3006	0.0079	0.0143	0.0261	0.3118
Adjusted R^2	-0.0004	0.3000	0.0071	0.0134	0.0253	0.3104
Residual Std. Error	0.1953 (df=2379)	0.1634 (df=2378)	0.1946 (df=2378)	0.1940 (df=2378)	0.1928 (df=2378)	0.1622 (df=2375)
F Statistic	0.1234 (df=1; 2379)	641.0058*** (df=2; 2378)	9.0977*** (df=2; 2378)	18.5245*** (df=2; 2378)	21.3504*** (df=2; 2378)	296.4169*** (df=5; 2375)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.13. RQ 2.4 - Event Studies: Model 1-6

						Dependent variable: Odd_home
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	0.0048	-0.0025	0.0035	0.0136	0.0019	0.0039
	(0.0104)	(0.0108)	(0.0107)	(0.0101)	(0.0111)	(0.0109)
C(Year)[T.2016]	0.0055	0.0041	0.0018	0.0121	0.0023	0.0075
	(0.0116)	(0.0125)	(0.0119)	(0.0116)	(0.0119)	(0.0126)
C(Year)[T.2017]	0.0018	-0.0041	0.0017	0.0066	0.0019	-0.0001
	(0.0116)	(0.0109)	(0.0116)	(0.0115)	(0.0118)	(0.0108)
C(Year)[T.2018]	0.0126	0.0035	0.0091	-0.0044	0.0072	-0.0122
	(0.0107)	(0.0112)	(0.0109)	(0.0110)	(0.0113)	(0.0114)
C(Year)[T.2019]	0.0195^{*}	0.0163	0.0161	0.0161	0.0155	0.0116
	(0.0118)	(0.0115)	(0.0118)	(0.0117)	(0.0119)	(0.0112)
C(Year)[T.2020]	0.0050	0.0103	0.0037	-0.0004	0.0059	0.0053
	(0.0115)	(0.0119)	(0.0117)	(0.0114)	(0.0119)	(0.0118)
C(Year)[T.2021]	-0.0166^{*}	-0.0065	-0.0170^{*}	-0.0243^{**}	-0.0162	-0.0132
	(0.0100)	(0.0111)	(0.0102)	(0.0099)	(0.0107)	(0.0109)
HomeAdvantage		0.0602^{***}				0.0589^{***}
		(0.0017)				(0.0016)
Intercept	0.4281^{***}	0.4108^{***}	0.4032^{***}	0.1621^{***}	0.3949^{***}	0.1810^{***}
	(0.0082)	(0.0084)	(0.0099)	(0.0440)	(0.0099)	(0.0367)
PlayingTime				0.0051^{***}		0.0041^{***}
				(0.0008)		(0.0007)
TotalGoals			0.0105^{***}			0.0045^{*}
			(0.0026)			(0.0024)
WinnerMargin					0.0260^{***}	0.0031
					(0.0040)	(0.0030)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0026	0.3017	0.0100	0.0174	0.0280	0.3135
Adjusted R^2	-0.0004	0.2993	0.0067	0.0140	0.0247	0.3103
Residual Std. Error	0.1953 (df=2373)	0.1635 (df=2372)	0.1946 (df=2372)	0.1939 (df=2372)	0.1929 (df=2372)	0.1622 (df=2369)
F Statistic	2.4782^{**} (df=7; 2373)	166.6047^{***} (df=8; 2372)	4.2511^{***} (df=8; 2372)	6.7974^{***} (df=8; 2372)	7.0636^{***} (df=8; 2372)	140.1360^{***} (df=11; 2369)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.14. RQ 2.4 - Event Studies: Models 7-12

	Dependent variable: Odd_home
	Model 13
HomeAdvantage	0.0593***
	(0.0023)
Intercept	0.1518***
	(0.0478)
PlayingTime	0.0047***
	(0.0009)
TotalGoals	0.0056*
	(0.0033)
Treatment	0.0816
	(0.0693)
Treatment:HomeAdvantage	-0.0008
	(0.0033)
Treatment:PlayingTime	-0.0016
	(0.0013)
Treatment:TotalGoals	-0.0018
	(0.0048)
Treatment:WinnerMargin	0.0013
	(0.0060)
WinnerMargin	0.0025
	(0.0043)
N	2381.0
Observations	2381
R^2	0.3122
Adjusted R^2	0.3096
Residual Std. Error	0.1623 (df=2371)
F Statistic	167.2829*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.15. RQ 2.4 - Event Studies: Model 13

						Dependent variable: PlayingTime
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	51.0131***	50.1422***	50.9687***	51.4774***	50.6519***	50.5257***
	(0.1479)	(0.1949)	(0.1474)	(0.2198)	(0.1780)	(0.2551)
Treatment	2.6947***	2.6608***	2.7043***	2.7097***	2.6898***	2.6750***
	(0.2086)	(0.2086)	(0.2084)	(0.2080)	(0.2088)	(0.2076)
Odd_away		0.1984***				0.2109***
		(0.0290)				(0.0331)
HomeAdvantage			0.1262**			-0.0971*
			(0.0539)			(0.0587)
TotalGoals				-0.1860***		-0.3891***
				(0.0593)		(0.0660)
WinnerMargin					0.2697***	0.4237***
					(0.0836)	(0.0937)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0752	0.0935	0.0773	0.0789	0.0795	0.1076
Adjusted R^2	0.0748	0.0927	0.0765	0.0781	0.0788	0.1058
Residual Std. Error	4.7253 (df=2379)	4.6792 (df=2378)	4.7209 (df=2378)	4.7168 (df=2378)	4.7151 (df=2378)	4.6456 (df=2375)
F Statistic	166.8638*** (df=1; 2379)	109.5065*** (df=2; 2378)	87.6564*** (df=2; 2378)	92.3731*** (df=2; 2378)	91.2682*** (df=2; 2378)	59.2297*** (df=5; 2375)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.16. RQ 3.1 - Event Studies: Models 1-6

						Dependent variable: PlayingTime
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	-1.7307***	-1.7551***	-1.7451***	-1.7078***	-1.7594***	-1.7404***
	(0.4015)	(0.3949)	(0.3996)	(0.4007)	(0.3975)	(0.3886)
C(Year)[T.2016]	-1.2984***	-1.2971***	-1.3011***	-1.2340***	-1.3305***	-1.2115***
	(0.4042)	(0.4029)	(0.4049)	(0.4013)	(0.4027)	(0.3960)
C(Year)[T.2017]	-0.9409**	-0.9264**	-0.9525**	-0.9394**	-0.9397**	-0.9084**
	(0.3887)	(0.3848)	(0.3889)	(0.3895)	(0.3871)	(0.3824)
C(Year)[T.2018]	3.3258***	3.3199***	3.3079***	3.3871***	3.2723***	3.3803***
	(0.3647)	(0.3601)	(0.3631)	(0.3610)	(0.3674)	(0.3585)
C(Year)[T.2019]	0.6756*	0.5541	0.6692*	0.7373*	0.6346	0.6130
	(0.3920)	(0.3863)	(0.3935)	(0.3903)	(0.3911)	(0.3779)
C(Year)[T.2020]	1.0685***	0.9787***	1.0790***	1.0914***	1.0774***	1.0212***
	(0.3433)	(0.3387)	(0.3448)	(0.3408)	(0.3441)	(0.3326)
C(Year)[T.2021]	1.5093***	1.5746***	1.5290***	1.5175***	1.5131***	1.5837***
	(0.3952)	(0.3900)	(0.3958)	(0.3932)	(0.3960)	(0.3854)
HomeAdvantage			0.1185**			-0.1180**
			(0.0535)			(0.0582)
Intercept	52.0621***	51.1477***	52.0281***	52.5046***	51.7278***	51.4913***
	(0.2991)	(0.3337)	(0.3001)	(0.3423)	(0.3201)	(0.3717)
Odd_away		0.2089***				0.2276***
		(0.0286)				(0.0327)
TotalGoals				-0.1866***		-0.3846***
				(0.0586)		(0.0649)
WinnerMargin					0.2614***	0.4065***
					(0.0827)	(0.0921)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.1046	0.1247	0.1064	0.1083	0.1086	0.1387
Adjusted R^2	0.1019	0.1218	0.1034	0.1053	0.1056	0.1347
Residual Std. Error	4.6555 (df=2373)	4.6037 (df=2372)	4.6517 (df=2372)	4.6468 (df=2372)	4.6459 (df=2372)	4.5698 (df=2369)
F Statistic	51.0484*** (df=7; 2373)	53.6251*** (df=8; 2372)	46.7597*** (df=8; 2372)	46.5219*** (df=8; 2372)	46.4448*** (df=8; 2372)	44.5767*** (df=11; 2369)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.17. RQ 3.1 - Event Studies: Models 7-12

	Dependent variable: PlayingTime
	Model 13
HomeAdvantage	-0.0992
	(0.0952)
Intercept	50.1149***
	(0.4128)
Odd_away	0.3367***
	(0.0692)
TotalGoals	-0.4046***
	(0.1035)
Treatment	3.2964***
	(0.5062)
Treatment:HomeAdvantage	-0.0203
	(0.1223)
Treatment:Odd_away	-0.1928**
	(0.0792)
Treatment:TotalGoals	0.0260
	(0.1341)
Treatment:WinnerMargin	0.1305
	(0.1857)
WinnerMargin	0.3476***
	(0.1310)
N	2381.0
Observations	2381
R^2	0.1116
Adjusted R^2	0.1083
Residual Std. Error	4.6390 (df=2371)
F Statistic	33.6808*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.18. RQ 3.1 - Event Studies: Model 13

	Dependent variable: Percentage_PlayingTime					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	0.5341***	0.5210***	0.5335***	0.5378***	0.5268***	0.5232***
	(0.0016)	(0.0022)	(0.0016)	(0.0025)	(0.0020)	(0.0029)
Treatment	0.0075***	0.0069***	0.0076***	0.0076***	0.0074***	0.0070***
	(0.0023)	(0.0022)	(0.0023)	(0.0022)	(0.0023)	(0.0022)
HomeAdvantage			0.0016***			-0.0019***
			(0.0006)			(0.0007)
Odd_away		0.0030***				0.0032***
		(0.0003)				(0.0004)
TotalGoals				-0.0015**		-0.0048***
				(0.0007)		(0.0008)
WinnerMargin					0.0054***	0.0072***
					(0.0010)	(0.0011)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0051	0.0416	0.0082	0.0071	0.0206	0.0672
Adjusted R^2	0.0047	0.0408	0.0074	0.0063	0.0198	0.0652
Residual Std. Error	0.0522 (df=2379)	0.0513 (df=2378)	0.0522 (df=2378)	0.0522 (df=2378)	0.0518 (df=2378)	0.0506 (df=2375)
F Statistic	10.9808*** (df=1; 2379)	52.0781*** (df=2; 2378)	10.4305*** (df=2; 2378)	7.8166*** (df=2; 2378)	23.0151*** (df=2; 2378)	35.7599*** (df=5; 2375)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.19. RQ 3.2 - Event Studies: Models 1-6

	Dependent variable: Percentage_PlayingTime					
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
C(Year)[T.2015]	-0.0200***	-0.0204***	-0.0202***	-0.0199***	-0.0206***	-0.0204***
	(0.0044)	(0.0044)	(0.0044)	(0.0044)	(0.0043)	(0.0042)
C(Year)[T.2016]	-0.0143***	-0.0143***	-0.0143***	-0.0139***	-0.0150***	-0.0136***
	(0.0044)	(0.0044)	(0.0045)	(0.0044)	(0.0044)	(0.0043)
C(Year)[T.2017]	-0.0106**	-0.0103**	-0.0107**	-0.0105**	-0.0105**	-0.0101**
	(0.0043)	(0.0042)	(0.0043)	(0.0043)	(0.0042)	(0.0042)
C(Year)[T.2018]	-0.0179***	-0.0180***	-0.0182***	-0.0175***	-0.0191***	-0.0178***
	(0.0043)	(0.0042)	(0.0043)	(0.0043)	(0.0043)	(0.0042)
C(Year)[T.2019]	-0.0036	-0.0053	-0.0037	-0.0032	-0.0045	-0.0049
	(0.0045)	(0.0044)	(0.0045)	(0.0044)	(0.0044)	(0.0042)
C(Year)[T.2020]	-0.0020	-0.0033	-0.0018	-0.0018	-0.0018	-0.0027
	(0.0039)	(0.0038)	(0.0039)	(0.0039)	(0.0039)	(0.0038)
C(Year)[T.2021]	0.0059	0.0068	0.0062	0.0059	0.0060	0.0069*
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0041)
HomeAdvantage			0.0018***			-0.0017***
			(0.0006)			(0.0006)
Intercept	0.5459***	0.5326***	0.5454***	0.5488***	0.5385***	0.5342***
	(0.0033)	(0.0037)	(0.0033)	(0.0038)	(0.0036)	(0.0042)
Odd_away		0.0030***				0.0031***
		(0.0003)				(0.0004)
TotalGoals				-0.0012*		-0.0046***
				(0.0007)		(0.0008)
WinnerMargin					0.0058***	0.0074***
					(0.0010)	(0.0010)
N	2381.0	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381	2381
R^2	0.0276	0.0652	0.0315	0.0290	0.0453	0.0904
Adjusted R^2	0.0247	0.0621	0.0282	0.0257	0.0421	0.0862
Residual Std. Error	0.0517 (df=2373)	0.0507 (df=2372)	0.0516 (df=2372)	0.0517 (df=2372)	0.0513 (df=2372)	0.0501 (df=2369)
F Statistic	11.1788*** (df=7; 2373)	22.7847*** (df=8; 2372)	12.2906*** (df=8; 2372)	10.0322*** (df=8; 2372)	15.9539*** (df=8; 2372)	22.2147*** (df=11; 2369)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.20. RQ 3.2 - Event Studies: Models 7-12

	Dependent variable: Percentage_PlayingTime
	Model 13
HomeAdvantage	-0.0012
	(0.0010)
Intercept	0.5200***
	(0.0044)
Odd_away	0.0038***
	(0.0007)
TotalGoals	-0.0040***
	(0.0011)
Treatment	0.0122**
	(0.0057)
Treatment:HomeAdvantage	-0.0014
	(0.0013)
Treatment:Odd_away	-0.0010
	(0.0009)
Treatment:TotalGoals	-0.0015
	(0.0015)
Treatment:WinnerMargin	0.0024
	(0.0021)
WinnerMargin	0.0058***
	(0.0015)
N	2381.0
Observations	2381
R^2	0.0695
Adjusted R^2	0.0660
Residual Std. Error	0.0506 (df=2371)
F Statistic	19.8071*** (df=9; 2371)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.21. RQ 3.2 - Event Studies: Model 13

					Dependent variable: WinnerMargin
	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	1.3394***	0.9707***	0.9338***	0.4072***	0.0871*
	(0.0328)	(0.0512)	(0.0358)	(0.0424)	(0.0514)
Treatment	0.0179	0.0036	0.0104	-0.0122	-0.0189
	(0.0498)	(0.0495)	(0.0464)	(0.0408)	(0.0381)
HomeWin			0.9137***		0.7082***
			(0.0429)		(0.0374)
Odd_away		0.0840***			0.0212***
		(0.0099)			(0.0073)
TotalGoals				0.3734***	0.3384***
				(0.0178)	(0.0171)
N	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381
R^2	0.0001	0.0550	0.1433	0.2502	0.3517
Adjusted R^2	-0.0004	0.0542	0.1426	0.2496	0.3506
Residual Std. Error	1.2008 (df=2379)	1.1676 (df=2378)	1.1117 (df=2378)	1.0400 (df=2378)	0.9675 (df=2376)
F Statistic	0.1298 (df=1; 2379)	36.1352*** (df=2; 2378)	227.7397*** (df=2; 2378)	223.5531*** (df=2; 2378)	273.8644*** (df=4; 2376)
Note:					*p<0.1; **p<0.05; ***p<0.01

Table B.22. RQ 4 - Event Studies: Models 1-5

					Dependent variable: WinnerMargin
	Model 6	Model 7	Model 8	Model 9	Model 10
C(Year)[T.2015]	0.1097	0.0999	0.1118	0.0641	0.0674
	(0.1046)	(0.1071)	(0.0976)	(0.0836)	(0.0795)
C(Year)[T.2016]	0.1228	0.1233	0.1368	-0.0055	0.0173
	(0.0943)	(0.0948)	(0.0888)	(0.0748)	(0.0700)
C(Year)[T.2017]	-0.0047	0.0012	0.0004	-0.0077	-0.0020
	(0.0912)	(0.0913)	(0.0841)	(0.0740)	(0.0696)
C(Year)[T.2018]	0.2045*	0.2021*	0.1442	0.0822	0.0462
	(0.1093)	(0.1099)	(0.1029)	(0.0904)	(0.0851)
C(Year)[T.2019]	0.1569	0.1080	0.1427	0.0338	0.0217
	(0.1041)	(0.1017)	(0.0977)	(0.0823)	(0.0764)
C(Year)[T.2020]	-0.0341	-0.0702	0.0007	-0.0797	-0.0578
	(0.0972)	(0.0977)	(0.0914)	(0.0762)	(0.0729)
C(Year)[T.2021]	-0.0145	0.0118	0.0173	-0.0309	0.0020
	(0.1046)	(0.1069)	(0.0977)	(0.0867)	(0.0824)
HomeWin			0.9086***		0.7047***
			(0.0432)		(0.0374)
Intercept	1.2792***	0.9112***	0.8703***	0.3969***	0.0673
	(0.0746)	(0.0855)	(0.0725)	(0.0678)	(0.0725)
Odd_away		0.0840***			0.0215***
		(0.0098)			(0.0073)
TotalGoals				0.3721***	0.3376***
				(0.0178)	(0.0171)
N	2381.0	2381.0	2381.0	2381.0	2381.0
Observations	2381	2381	2381	2381	2381
R^2	0.0051	0.0598	0.1463	0.2519	0.3525
Adjusted R^2	0.0022	0.0566	0.1434	0.2494	0.3498
Residual Std. Error	1.1993 (df=2373)	1.1661 (df=2372)	1.1112 (df=2372)	1.0402 (df=2372)	0.9681 (df=2370)
F Statistic	1.7493* (df=7; 2373)	11.1264*** (df=8; 2372)	58.0149*** (df=8; 2372)	56.6449*** (df=8; 2372)	111.9332*** (df=10; 2370)
Note:					*p<0.1; **p<0.05; ***p<0.01

Table B.23. RQ 4 - Event Studies: Models 6-10

	Dependent variable: WinnerMargin
	Model 11
HomeWin	0.7805***
	(0.0517)
Intercept	0.0434
	(0.0616)
Odd_away	0.0246**
	(0.0099)
TotalGoals	0.3370***
	(0.0227)
Treatment	0.0624
	(0.0914)
Treatment:HomeWin	-0.1448**
	(0.0735)
Treatment:Odd_away	-0.0043
	(0.0141)
Treatment:TotalGoals	0.0011
	(0.0338)
N	2381.0
Observations	2381
R^2	0.3528
Adjusted R^2	0.3509
Residual Std. Error	0.9673 (df=2373)
F Statistic	161.3974*** (df=7; 2373)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table B.24. RQ 4 - Event Studies: Model 11

Appendix C- Generalized Linear Models

	Dependent variable: TotalGoals
	Model GLM
HomeWin	-0.0498
	(0.0311)
Intercept	1.2176***
	(0.1399)
Odd_home	0.1118
	(0.0736)
PlayingTime	-0.0131***
	(0.0027)
Treatment	0.0609**
	(0.0268)
WinnerMargin	0.2249***
	(0.0093)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: TotalGoals
	Model GLM
C(Year)[T.2015]	-0.0128
	(0.0557)
C(Year)[T.2016]	0.0719
	(0.0547)
C(Year)[T.2017]	-0.0105
	(0.0560)
C(Year)[T.2018]	0.1164**
	(0.0554)
C(Year)[T.2019]	0.0824
	(0.0546)
C(Year)[T.2020]	0.0608
	(0.0555)
C(Year)[T.2021]	0.0397
	(0.0560)
HomeWin	-0.0500
	(0.0312)
Intercept	1.2210***
	(0.1497)
Odd_home	0.1090
	(0.0738)
PlayingTime	-0.0134***
	(0.0028)
WinnerMargin	0.2234***
	(0.0093)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

Table C.1. RQ 1 - GLM simple and year fixed effect

	Dependent variable: HomeAdvantage
	Model GLM
Intercept	1.6311***
	(0.3473)
Odd_away	-5.0478***
	(0.1699)
PlayingTime	-0.0002
	(0.0065)
TotalGoals	-0.0301
	(0.0220)
Treatment	-0.0810
	(0.0633)
WinnerMargin	0.2910***
	(0.0295)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: HomeAdvantage
	Model GLM
C(Year)[T.2015]	0.0621
	(0.1284)
C(Year)[T.2016]	-0.0265
	(0.1283)
C(Year)[T.2017]	0.0610
	(0.1280)
C(Year)[T.2018]	0.0621
	(0.1299)
C(Year)[T.2019]	-0.0622
	(0.1282)
C(Year)[T.2020]	-0.0958
	(0.1281)
C(Year)[T.2021]	-0.1134
	(0.1283)
Intercept	1.6667***
	(0.3694)
Odd_away	-5.0448***
	(0.1701)
PlayingTime	-0.0014
	(0.0066)
TotalGoals	-0.0305
	(0.0220)
WinnerMargin	0.2896***
	(0.0295)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

Table C.2. RQ 2.1 - GLM simple and year fixed effect

	Dependent variable: HomeWin
	Model GLM
Intercept	1.3323^{**}
	(0.5728)
Odd_away	-7.0100^{***}
	(0.3768)
PlayingTime	-0.0093
	(0.0107)
TotalGoals	-0.1211^{***}
	(0.0364)
Treatment	0.0606
	(0.1049)
WinnerMargin	1.0060^{***}
	(0.0593)
N	2381
Observations	2381
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

	Dependent variable: HomeWin
	Model GLM
C(Year)[T.2015]	-0.1182
	(0.2153)
C(Year)[T.2016]	-0.2022
	(0.2143)
C(Year)[T.2017]	-0.0668
	(0.2115)
C(Year)[T.2018]	0.2203
	(0.2155)
C(Year)[T.2019]	-0.0645
	(0.2118)
C(Year)[T.2020]	-0.1663
	(0.2136)
C(Year)[T.2021]	-0.1065
	(0.2109)
Intercept	1.6426^{***}
	(0.6126)
Odd_away	-7.0217^{***}
	(0.3775)
PlayingTime	-0.0133
	(0.0109)
TotalGoals	-0.1223^{***}
	(0.0366)
WinnerMargin	1.0068^{***}
	(0.0595)
N	2381
Observations	2381
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table C.3. RQ 2.2 - GLM simple and year fixed effect

	Dependent variable: HomePoints
	Model GLM
Intercept	-0.2579
	(0.2991)
Odd_home	2.1214***
	(0.1412)
PlayingTime	-0.0059
	(0.0058)
TotalGoals	-0.0347*
	(0.0195)
Treatment	0.0062
	(0.0557)
WinnerMargin	0.0790***
	(0.0259)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: HomePoints
	Model GLM
C(Year)[T.2015]	-0.0392
	(0.1126)
C(Year)[T.2016]	-0.0731
	(0.1129)
C(Year)[T.2017]	-0.0205
	(0.1121)
C(Year)[T.2018]	0.0669
	(0.1130)
C(Year)[T.2019]	-0.0616
	(0.1125)
C(Year)[T.2020]	-0.0749
	(0.1129)
C(Year)[T.2021]	-0.0343
	(0.1132)
Intercept	-0.1429
	(0.3183)
Odd_home	2.1286***
	(0.1416)
PlayingTime	-0.0075
	(0.0059)
TotalGoals	-0.0348*
	(0.0196)
WinnerMargin	0.0785***
	(0.0259)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

Table C.4. RQ 2.3 - GLM simple and year fixed effect

	Dependent variable: Odd_home
	Model GLM
HomeAdvantage	0.0589***
	(0.0019)
Intercept	0.1947***
	(0.0371)
PlayingTime	0.0039***
	(0.0007)
TotalGoals	0.0046*
	(0.0024)
Treatment	-0.0045
	(0.0069)
WinnerMargin	0.0032
	(0.0033)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: Odd_home
	Model GLM
C(Year)[T.2015]	0.0039
	(0.0140)
C(Year)[T.2016]	0.0075
	(0.0140)
C(Year)[T.2017]	-0.0001
	(0.0140)
C(Year)[T.2018]	-0.0122
	(0.0142)
C(Year)[T.2019]	0.0116
	(0.0140)
C(Year)[T.2020]	0.0053
	(0.0140)
C(Year)[T.2021]	-0.0132
	(0.0140)
HomeAdvantage	0.0589***
	(0.0019)
Intercept	0.1810***
	(0.0395)
PlayingTime	0.0041***
	(0.0007)
TotalGoals	0.0045*
	(0.0024)
WinnerMargin	0.0031
	(0.0033)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

Table C.5. RQ 2.4 - GLM simple and year fixed effect

	Dependent variable: PlayingTime
	Model GLM
HomeAdvantage	-0.0971
	(0.0617)
Intercept	50.5257***
	(0.2349)
Odd_away	0.2109***
	(0.0329)
TotalGoals	-0.3891***
	(0.0684)
Treatment	2.6750***
	(0.1907)
WinnerMargin	0.4237***
	(0.0942)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: PlayingTime
	Model GLM
C(Year)[T.2015]	-1.7404***
	(0.3942)
C(Year)[T.2016]	-1.2115***
	(0.3946)
C(Year)[T.2017]	-0.9084**
	(0.3941)
C(Year)[T.2018]	3.3803***
	(0.3946)
C(Year)[T.2019]	0.6130
	(0.3951)
C(Year)[T.2020]	1.0212***
	(0.3945)
C(Year)[T.2021]	1.5837***
	(0.3942)
HomeAdvantage	-0.1180*
	(0.0608)
Intercept	51.4913***
	(0.3479)
Odd_away	0.2276***
	(0.0324)
TotalGoals	-0.3846***
	(0.0674)
WinnerMargin	0.4065***
	(0.0928)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

Table C.6. RQ 3.1 - GLM simple and year fixed effect

	Dependent variable: Percentage_PlayingTime
	Model GLM
HomeAdvantage	-0.0019***
	(0.0007)
Intercept	0.5232***
	(0.0026)
Odd_away	0.0032***
	(0.0004)
TotalGoals	-0.0048***
	(0.0007)
Treatment	0.0070***
	(0.0021)
WinnerMargin	0.0072***
	(0.0010)
N	2381
Observations	2381
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: Percentage_PlayingTime
	Model GLM
C(Year)[T.2015]	-0.0204***
	(0.0043)
C(Year)[T.2016]	-0.0136***
	(0.0043)
C(Year)[T.2017]	-0.0101**
	(0.0043)
C(Year)[T.2018]	-0.0178***
	(0.0043)
C(Year)[T.2019]	-0.0049
	(0.0043)
C(Year)[T.2020]	-0.0027
	(0.0043)
C(Year)[T.2021]	0.0069
	(0.0043)
HomeAdvantage	-0.0017***
	(0.0007)
Intercept	0.5342***
	(0.0038)
Odd_away	0.0031***
	(0.0004)
TotalGoals	-0.0046***
	(0.0007)
WinnerMargin	0.0074***
	(0.0010)
N	2381
Observations	2381

Table C.7. RQ 3.2 - GLM simple and year fixed effect

	Dependent variable: WinnerMargin
	Model GLM
HomeWin	0.5673***
	(0.0396)
Intercept	-0.6830***
	(0.0461)
Odd_away	0.0066
	(0.0050)
TotalGoals	0.2276***
	(0.0101)
Treatment	-0.0172
	(0.0354)
inflate_const	-18.2535
	(304.5604)
N	2381
Observations	2381
Residual Std. Error	0.9822 (df=2376)
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: WinnerMargin
	Model GLM
C(Year)[T.2015]	0.0379
	(0.0749)
C(Year)[T.2016]	-0.0085
	(0.0749)
C(Year)[T.2017]	-0.0321
	(0.0763)
C(Year)[T.2018]	0.0090
	(0.0741)
C(Year)[T.2019]	-0.0284
	(0.0750)
C(Year)[T.2020]	-0.0595
	(0.0769)
C(Year)[T.2021]	-0.0054
	(0.0764)
HomeWin	0.5650***
	(0.0398)
Intercept	-0.6818***
	(0.0678)
Odd_away	0.0070
	(0.0051)
TotalGoals	0.2276***
	(0.0102)
inflate_const	-11.6465
	(11.1854)
N	2381
Observations	2381
Residual Std. Error	0.9821 (df=2370)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table C.8. RQ 4 - GLM simple and year fixed effect

Appendix D- Regression Discontinuity Design

	Dependent variable: TotalGoals
	Model 1
HomeWin	-0.2462***
	(0.0873)
Intercept	3.0420***
	(0.3988)
Odd_home	0.4979**
	(0.2090)
PlayingTime	-0.0335***
	(0.0078)
Running_Var	-0.0991*
	(0.0520)
Treatment	0.3920**
	(0.1536)
WinnerMargin	0.7142***
	(0.0323)
N	1529.0
Observations	1529
R^2	0.2683
Adjusted R^2	0.2655
Residual Std. Error	1.4232 (df=1522)
F Statistic	93.0365*** (df=6; 1522)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D.1. RQ 1 - Regression Discontinuity Design

	Dependent variable: HomeAdvantage
	Model 1
Intercept	1.4208***
	(0.4355)
Odd_away	-5.0374***
	(0.2139)
PlayingTime	0.0013
	(0.0082)
Running_Var	-0.0389
	(0.0548)
TotalGoals	-0.0156
	(0.0269)
Treatment	0.0383
	(0.1620)
WinnerMargin	0.3080***
	(0.0367)
N	1529.0
Observations	1529
R^2	0.3123
Adjusted R^2	0.3096
Residual Std. Error	1.4995 (df=1522)
F Statistic	115.2163*** (df=6; 1522)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D.2. RQ 2.1 - Regression Discontinuity Design

	Dependent variable: HomeWin	
	Model 1	Model 2
Intercept	-0.4847***	-0.1811
	(0.1365)	(1981734.6952)
Running_Minus_Cutoff	-0.1609**	0.0431
	(0.0728)	(2972602.0429)
Running_Minus_Cutoff_Squared		0.0011
		(990867.3476)
Treatment	0.5030**	0.2465
	(0.2099)	(1981734.6952)
Treatment_x_Running		-0.2387
		(2972602.0429)
Treatment_x_Running_Squared		-0.0088
		(990867.3476)
N	1529	1529
Observations	1529	1529
Note:		*p<0.1; **p<0.05; ***p<0.01

Table D.3. RQ 2.2 - Regression Discontinuity Designs

	Dependent variable: HomePoints
	Model 1
Intercept	0.2299
	(0.3279)
Odd_home	3.1407^{***}
	(0.1528)
PlayingTime	-0.0051
	(0.0063)
Running_Var	-0.0701^{*}
	(0.0420)
TotalGoals	-0.0354^{*}
	(0.0206)
Treatment	0.1968
	(0.1241)
WinnerMargin	0.1753^{***}
	(0.0283)
N	1529.0
Observations	1529
R^2	0.2594
Adjusted R^2	0.2565
Residual Std. Error	1.1482 (df=1522)
F Statistic	88.8699^{***} (df=6; 1522)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table D.4. RQ 2.3 - Regression Discontinuity Design

	Dependent variable: Odd_home
	Model 1
HomeAdvantage	0.0586***
	(0.0024)
Intercept	0.2334***
	(0.0464)
PlayingTime	0.0033***
	(0.0009)
Running_Var	0.0049
	(0.0060)
TotalGoals	0.0047
	(0.0029)
Treatment	-0.0122
	(0.0176)
WinnerMargin	0.0045
	(0.0041)
N	1529.0
Observations	1529
R^2	0.3151
Adjusted R^2	0.3124
Residual Std. Error	0.1631 (df=1522)
F Statistic	116.7019*** (df=6; 1522)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D.5. RQ 2.4 - Regression Discontinuity Design

	Dependent variable: PlayingTime
	Model 1
HomeAdvantage	-0.0862
	(0.0766)
Intercept	49.2314***
	(0.3881)
Odd_away	0.1758***
	(0.0400)
Running_Var	-0.8652***
	(0.1686)
TotalGoals	-0.3493***
	(0.0829)
Treatment	4.9123***
	(0.4860)
WinnerMargin	0.4258***
	(0.1164)
N	1529.0
Observations	1529
R^2	0.1191
Adjusted R^2	0.1156
Residual Std. Error	4.6400 (df=1522)
F Statistic	34.2836*** (df=6; 1522)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D.6. RQ 3.1 - Regression Discontinuity Design

	Dependent variable: Percentage_PlayingTime
	Model 1
HomeAdvantage	-0.0019***
	(0.0007)
Intercept	0.5306***
	(0.0034)
Odd_away	0.0032***
	(0.0004)
Running_Var	0.0032***
	(0.0009)
TotalGoals	-0.0047***
	(0.0007)
Treatment	-0.0054
	(0.0042)
WinnerMargin	0.0073***
	(0.0010)
N	2381.0
Observations	2381
R^2	0.0716
Adjusted R^2	0.0693
Residual Std. Error	0.0505 (df=2374)
F Statistic	30.5165*** (df=6; 2374)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D.7. RQ 3.2 - Regression Discontinuity Design

	Dependent variable: WinnerMargin
	Model 1
HomeWin	0.6668***
	(0.0559)
Intercept	-0.0090
	(0.0805)
Odd_away	0.0316***
	(0.0080)
Running_Var	-0.0495
	(0.0358)
TotalGoals	0.3309***
	(0.0154)
Treatment	0.1171
	(0.1031)
N	1529.0
Observations	1529
R^2	0.3548
Adjusted R^2	0.3527
Residual Std. Error	0.9840 (df=1523)
F Statistic	167.5177*** (df=5; 1523)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D.8. RQ 4 - Regression Discontinuity Design

NOVA Information Management School

Instituto Superior de Estatística e Gestão de Empresas

Universidade Nova de Lisboa