Essays on Behavioural Economics of Sport

KOCSOY, ALPEREN

Essays on

Behavioural Economics of Sport

Alperen Kocsoy

Department of Economics

Durham University

This thesis is submitted for the degree of

Doctor of Philosophy

Ustinov College September 2024

To the hero stating that the truest guide in life is science

and

To the young dreamer lost in the literature class in MCMV...

Declaration

I, Alperen Kocsoy, hereby declare that this thesis, titled "Essays on Behavioural Economics of Sport", is entirely my own work. All sources and materials used in the preparation of this thesis have been properly acknowledged and cited. I confirm that no part of this work has been submitted for any other degree or qualification at this or any other university or institution.

Alperen Kocsoy

September 2024

Statement of Copyright

Acknowledgements

Completing this thesis has been a challenging yet rewarding journey, one that would not have been possible without the support, guidance, and contributions of numerous people. I would like to express my sincere gratitude to everyone who has played a role in helping me reach this milestone.

I would like to express my profound gratitude to the Republic of Türkiye, 'my beloved, beautiful, yet solitary country'[1]. A scholarship program from Türkiye made my doctoral studies and this thesis possible.

Next, I would like to express my heartfelt gratitude to my supervisors, Nejat Anbarcı and Ángel Hernando-Veciana, for their guidance and mentorship throughout this journey. Your patience and dedication have been instrumental in the completion of this thesis. I am especially indebted to Nejat, without whom I would have been dropped in the very first year. Your support and belief in my abilities have been a driving force behind my perseverance and efforts.

I would like to thank Damian Damianov, Anil Yildizparlak, John Moffat, and Mehmet Ismail for their invaluable feedback throughout the revision and examination of this thesis. Your insights and constructive feedback have enhanced the quality of this work.

I extend my appreciation to Carl Singleton, Paul Downward, and Robert Butler for their thoughtful feedback. To Cem Cakmakli, Eren Arbatli, Ercument Aksak, and Ibrahim Inal, I am deeply grateful for your unwavering support and the valuable feedback you have provided. Your expertise and knowledge were invaluable in shaping my papers and refining my ideas.

[1]Nuri Bilge Ceylan, Cannes, 2008

I am also grateful to Baris Gerceker and Opta for generously providing a dataset that has been instrumental in my research.

I would also like to express my appreciation to all the other academics from the departments of Economics and Finance with whom I had the privilege of teaching. Your guidance and support have been invaluable in shaping my early career as an educator. To my adorable students, thank you for making my teaching journey an unforgettable and joyous experience.

I will always cherish the memories of our weekly gatherings at Duke of Wellington with the Turkish academics from the department. The laughter, stimulating conversations, and sense of community will be deeply missed.

I would also like to thank Frances, the best landlady and housemate one could ever ask for. Your kindness and support have made my time in Durham during this journey all the more enjoyable.

To Yunus, my best friend, I am forever grateful for your patience and understanding as you listened to all my crazy ideas. Your friendship and support have been a constant source of comfort and inspiration.

To the members of the Durham University Bridge Club and Durham Bridge Club, as well as all my partners, teammates, and competitors, thank you for the countless enjoyable evenings spent playing bridge. Your camaraderie and the friendly competition have been a welcome respite from the rigours of academic life.

Finally, I would like to thank my family for their unconditional love, support, and encouragement throughout this challenging journey. Your belief in me has been a constant source of strength and motivation, even though most of the time you had no idea what I was doing.

To everyone mentioned here, and to those whom I may have inadvertently omitted, please know that your contributions, no matter how big or small, have been appreciated. This thesis is a testament to the collective effort and support of each and every one of you. Thank you.

Abstract

This thesis comprises three independent studies that explore various aspects of human behaviour, using data from sports. The studies are linked by their focus on how social, cultural, and individual factors influence the behaviour and decision-making process of individuals.

The first chapter introduces a novel variable, ball-in-play time, to assess potential referee bias in football, particularly regarding the decision on additional time. By comparing matches played behind closed doors during the Covid-19 pandemic with those played in front of fans, the study finds that referees exhibit a bias towards home teams, but only in the presence of fans. This suggests that social pressure from fans significantly influences the decisions of referees. The study also examines referee behaviour when there is a strength difference between the competing teams to investigate 'big team' or 'favourite team' bias.

The second chapter investigates the impact of formally assigned leaders (team captains) and informal leaders (all-stars) on their teammates' productivity in basketball. Using in-game injuries as random shocks, the study employs a novel staggered difference-in-differences estimation to examine peer effects in high-stakes team environments. The key finding is that only players who are both formal and informal leaders have positive effects on their teammates' performance. These findings could extend to team management practices across various industries.

The third chapter explores the relationship between the patience levels of countries and the tenure time of managers, using data from football. By examining the tenure of football

managers across different countries, the study offers insights into employment practices in sports and provides a broader understanding of how cultural and socioeconomic factors shape employment practices in various sectors. The analysis shows a positive effect of a country's patience level on the tenure duration of football managers, which is confirmed through instrumental variables and survival analysis.

Together, these studies contribute to our understanding of how social pressure, leadership, and cultural factors influence decision-making in sports, with potential implications for other sectors.

Table of Contents

List of Tables

List of Figures

Chapter 1

Referee Bias in Football: Actual vs. Expected Additional Time

1.1 Introduction

Do social forces and pressure affect the decisions of individuals? Economists have long investigated how endogenous preferences are affected by such social environments (see, e.g. Akerlof, 1980, 1997; Akerlof & Kranton, 2000; Becker & Murphy, 2000). Bias, in its numerous forms, significantly influences human decision-making and plays an integral role in shaping outcomes across a broad range of economic contexts. The presence of bias can distort market efficiency, foster inequality, and lead to sub-optimal outcomes, causing a divergence from theoretical expectations based on models of rational behaviour (Arrow, 1951). The recognition of these impacts has stimulated extensive research within the field of economics to explore the existence, nature, and implications of bias (Becker, 1957).

On this matter, there is experimental and empirical evidence showing that individuals' behaviour is biased under different forms of social pressure. For example, Charness et al. (2007) and Charness & Sutter (2012) provide experimental evidence, while Garicano et al. (2005) and Sutter & Kocher (2004) offer empirical evidence from sports. Additionally, Dohmen & Sauermann (2016) present a literature survey on referee bias in sports.

Sports presents a suitable environment to explore biases. The socially charged context, where there is a passionate fandom, high-stakes competition, split-second decisions by referees and high emotions, creates suitable grounds for biases to manifest and influence decision-making.

The exploration of referee bias in football has been a focal area in this domain for more than two decades. Referees, under social pressure from fans, have been found to make decisions favouring the home teams (Nevill et al., 2002; Garicano et al., 2005). These findings lend experimental and empirical weight to the theory that social forces are able to bias individual decisions and help to understand bias in general economic contexts.

Previous research has proposed two main explanations for why referees may favour the home team. The first explanation, as argued by Nevill et al. (2002), is psychological in nature and suggests that crowd noise acts as a decision-making heuristic (Tversky & Kahneman, 1974). In this context, the likelihood of an incident being perceived as a foul is increased when accompanied by crowd noise, which tends to be louder as a reaction when a foul is committed by an away team player compared to a home team player.

Alternatively, Sutter & Kocher (2004) offers an economic perspective based on Agency Theory. They argue that referee bias can be seen as a rational, optimising response to the need for referees to balance appearing impartial to their employers (i.e., the football governing body) while also appeasing the crowd. In this scenario, referees are considered fully informed agents when making decisions. However, as monitoring by their employers is less complete than that of the immediate crowd, referees tend to show bias in favour of the home team (Downward & Jones, 2007).

These findings suggest that the presence of a crowd can influence the decisions of referees, either through the noise they generate or the referee's perception of being monitored. This bias may be a contributing factor to the well-documented phenomenon of home advantage in football (Downward & Jones, 2007; Nevill et al., 2002; Sutter & Kocher, 2004).

Football, also known as soccer in the new world, is a team sport played between two teams of 11 players each. The objective is to score goals by getting the ball into the opposing team's goal. A match consists of two 45-minute halves, with additional time (known as 'injury time' or 'stoppage time') added at the end of each half by a referee to compensate for delays during play. The match is officiated by a referee who is assisted by two assistant referees and a fourth official. The referee has the authority to enforce the rules of the game such as awarding free kicks for fouls, showing yellow cards (cautions) for serious fouls, and red cards for very serious offences or after a player receives two yellow cards which results in the player's dismissal from the game. The referee also decides how much additional time to add at the end of each half based on time lost due to substitutions, injuries, time-wasting or other stoppages. The team that scores more goals by the end of the match wins. If the scores are level at the end of 90^{th} minute, the match may end in a draw or, in some competitions, proceed to extra time or a penalty shootout.

The Covid-19 pandemic offered an unprecedented natural experiment to further investigate the effect of social pressure on referee decisions. Studies focusing on football matches played behind closed doors during the pandemic find that referees decided in favour of home teams only when fans were present (Bryson, Dolton, et al., 2021; Endrich & Gesche, 2020; Reade et al., 2022; Scoppa, 2021). This finding confirms the significant influence of social pressure on decision-making. These studies analyse the differences in referees' yellow and red card decisions. However, it is difficult to assess whether these decisions are just or unjust (Garicano et al., 2005) as both home and away teams may also adjust the toughness of their playing styles accordingly in the absence of fans, influencing the perceived home team advantage.

Moreover, while recognising the studies conducted on referee behaviour under social pressures, including fan presence, the literature is yet to fully explore how these dynamics have evolved post-Covid-19. While studies during the pandemic, with matches behind closed doors, find a potential reduction in referee bias, it remains unclear if these impartial decisions have continued as fans return to stadiums. This uncertainty presents a gap in the literature. There is a need for focused research to determine whether the return of fans has led to a re-emergence of referee bias or if the pandemic has had a lasting effect by fostering more unbiased refereeing.

Furthermore, other types of bias like racial and nationality biases have also been evidenced in sports. Studies have shown that referees favoured players of their own race and nationality (Price & Wolfers, 2010; Krumer et al., 2022). The public release of these findings, such as the study by Price & Wolfers (2010), has led to a reduction in racial bias, showing the role of awareness in mitigating such biases (D. G. Pope et al., 2018). This suggests that increasing transparency and promoting awareness are effective strategies for reducing referee bias, offering potential lessons for addressing other biases.

In football, referees have the discretion to decide on additional time to compensate for lost time during each half of the game (IFAB, 2022). This additional time, often called injury or stoppage time, can be very crucial and change the outcome of the game. Such moments not only decide matches but also create unforgettable events in football history.

On the other side, another form of bias in sports is the big team bias, where referees tend to favour teams with larger fan bases or historical success (Boyko et al., 2007; Lago-Peñas & Gómez-López, 2016; Bose et al., 2022). A well-known example in football is 'Fergie Time,' a term named after Sir Alex Ferguson, the former manager of Manchester United. 'Fergie Time' refers to the perceived extra additional time given at the end of games to benefit Ferguson's team, often seen as an advantage during close matches when a few more minutes could change the game's outcome (Butler & Butler, 2017). This phenomenon illustrates how big team bias can manifest as referees may subconsciously extend matches to favour more popular teams.

Although football analysts have long recognised the significance of goals scored during this period, often referred to as 'late goals', scientific evidence on the impact of additional time on the game's outcome is quite scarce. These goals can flip the momentum of the game, undoing the strategies and efforts put forth throughout the match. For instance, a team leading by one goal may suddenly find itself forced into a draw, or even worse, a loss due to a late goal in the additional time. Thus, additional time can either provide a lifeline for teams on the brink of defeat or deliver a fatal blow to those on the verge of victory.

Moreover, the tension and high stakes during additional time can exert significant pressure on players, potentially leading to critical mistakes or spectacular performances. It is not uncommon to observe more aggressive playing and tactical changes during this period, as teams scramble to change the scoreline before the final whistle. The Table 1.1 below illustrates the importance of the additional time on match outcome, using data from top European leagues between 2018 and 2023.

Table 1.1 *Changes in Match Outcome in the Additional Time (All Games)*

Before/After	Away Win	Draw	Home Win	Total
Away Win	94.54%	5.21%	0.25%	100%
Draw	5.75%	86.26%	7.98%	100%
Home Win	0.18%	3.97%	95.84%	100%
TOTAL	31.32%	25.63%	43.05%	100%

Note: The table shows proportional outcome changes from just before to after additional time. Data source: Opta.

When the 90 minutes end with a draw, a home team, which has a 7.98% chance, is around 39% more likely to win than an away team, which has a 5.75% chance. Although the chances of staging a comeback and winning the match may seem low, they could be higher when a team is only leading by one goal, as both teams are likely to take greater risks to score and secure points. In football, a win awards three points, a draw one point, and a loss zero points. Teams can make bold choices towards the end of a match because of loss aversion (Tversky & Kahneman, 1991) and the deadline day effect (Ariely & Wertenbroch, 2002). If the score is tied, some teams might risk the one point from a draw in an attempt to gain three points with a win. Similarly, if a team is behind, they use every second of additional time to try to score. Table 1.2 shows that a change in outcome is more likely when the score difference is at most one goal.

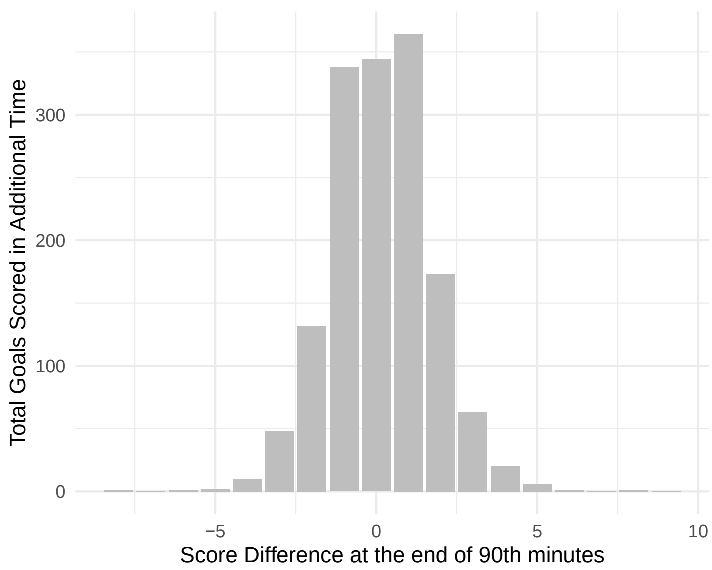
Table 1.2 *Changes in Match Outcome in the Additional Time (at Most 1-Goal Difference)*

Before/After	Away Win	Draw	Home Win	Total
Away Win	89.78%	9.75%	0.47%	100%
Draw	5.59%	86.56%	7.84%	100%
Home Win	0.40%	8.71%	90.88%	100%
TOTAL	26.56%	41.61%	31.83%	100%

Note: The table shows proportional outcome changes from just before to after additional time. Data source: Opta.

As illustrated in Figure 1.1, a significant proportion of goals scored during additional time occur when the match is either at a draw or when a team maintains a narrow lead. Consequently, decisions made to allocate additional time are of utmost importance; they can substantially influence the outcome of the match. This underscores the critical role that refereeing plays during these pivotal moments, emphasising the need for fair and impartial decision-making by referees.

Fig. 1.1 *Goals Scored in Additional Time and Score Difference (Home−Away) at the End of the 90th Minutes*



Additional times see more goals if the winner could be changed. Note: Data source: Opta.

Furthermore, additional time decisions may also have implications beyond the scope of a single match. These decisions can also affect a team's standing in the league, their promotion or relegation in round-robin tournaments or progression and elimination in knockout tournaments, and, indirectly, their financial status, given the significant financial rewards attached to successful competition performance. As such, the additional time becomes a microcosm of the broader competition, encapsulating the drama, unpredictability, and high stakes that make football a globally attractive sport.

The discretionary power entrusted to referees in determining additional time makes it an interesting subject for investigation. This autonomy can potentially introduce bias, leading to substantial variations in match outcomes. Consequently, it is crucial to examine the presence of referee bias in additional time decisions and explore potential solutions to mitigate such bias.

This study contributes to the literature by introducing a novel variable, ball-in-play time, to provide a more accurate estimation of referee bias in the allocation of additional time. Ball-in-play time represents the precise duration of active gameplay, accounting for every stoppage in the match. By incorporating this new variable, I aim to reduce potential omitted variable bias and yield more accurate estimates of referee bias. This methodological improvement not only advances the understanding of referee bias but also contributes to developing objective solutions for addressing it.

To systematically evaluate potential bias in referees' decisions regarding additional time, I employ a precise calculation method using ball-in-play data. According to this study and Garicano et al. (2005)[1], referee bias is non-existent in the first halves of football matches. During these periods, the ball is actively in play for approximately 62% of the total time. This statistic serves as a benchmark for computing the 'expected additional time' for each game, defined as the duration needed to ensure that active playing time constitutes 62% of the total time for a given half. I then compare the expected additional time with

[1]This could be because there are still 45 minutes to play in the second half, which makes the additional time in the first half seem less important (Garicano et al., 2005).

the actual additional time allocated by referees at the end of a match. A discrepancy dependent on whether the leading team is home or away is considered an indicator of 'bias'. By adapting and improving the empirical strategy of Garicano et al. (2005), I examine how the difference between expected and actual additional time at the end of the 90$^{\text{th}}$ minute, influenced by the presence of home fans, varies depending on whether the home or away team is leading by a one-goal margin.

While there was a positive and significant difference between the actual and expected additional time when home teams were leading versus losing both before and after the Covid-19 pandemic, this difference was not significant from the second quarter of 2020 to the same quarter of 2021, a period marked by the Covid-19 pandemic restrictions and the absence of fans. This suggests that the tendency of referees to add more time when home teams are trailing is notably influenced by the presence of fans.

Finally, I recommend a pragmatic solution to mitigate referee bias in the allocation of additional time. By employing ball-in-play data to calculate additional time automatically, we can introduce a fairer and more transparent system. This approach effectively reduces the subjective burden on referees by standardising the allocation process based on objective measures.

1.2 Literature Review

Bias in economic contexts can be understood as systematic deviations from the predictions of classical economic models, often due to individual preferences, cognitive limitations, or societal influences. It can alter the decision-making process and carries profound implications for economic behaviour and outcomes (Rabin, 1998; DellaVigna, 2009).

In this chapter, the literature review is divided into two subsections to effectively address the discussion of bias. The first part focuses on the broader concept of bias in economics, laying the groundwork to show how bias can impact decision-making in economics. The second part narrows our focus to bias within sports, offering a view of how bias manifests in sporting contests.

1.2.1 Bias in Economics

An early examination of bias in economics can be traced back to the works of Arrow (1951) and Becker (1957) on discrimination as a kind of bias in the labour markets. These studies laid the foundation for understanding how personal prejudices could distort market outcomes.

Building on this, a significant part of the literature has focused on cognitive biases in individual decision-making. Prospect theory of Kahneman & Tversky (1979) demonstrates how people often violate the assumptions of expected utility theory due to cognitive biases such as loss aversion and probability weighting. This seminal work has opened a road to several other studies which examine the consequences of cognitive biases in areas like consumer behaviour, investment decisions, and policy-making (Thaler, 2015).

Behavioural economics also emphasises the role of social influences in biased economic decision-making. The 'social distance' concept of Akerlof (1980) explains how preferences for association with similar individuals can cause discriminatory behaviour. This idea was further developed by Akerlof & Kranton (2000) in their 'identity economics' framework which highlights how social identity can bias the economic choices of individuals. Becker & Murphy (2000), in another seminal work, illustrates how social pressure could start 'rational addiction', by showing how harmful addictions might be a rational response to the influence of the social environment.

The role of bias also extends into public economics. Studies by Carozzi & Repetto (2016); Do et al. (2017) and Hodler & Raschky (2014) reveal evidence of political biases that are influencing government investment decisions by highlighting how the strategic allocation of public resources could be allocated by the biases of policymakers.

Bommer et al. (2022) exposes systematic preferences in international aid allocation by showing that resources are unevenly distributed due to geopolitical interests. Additionally, Kuziemko & Werker (2006) indicates that the history of colonial relationships can influence aid decisions, displaying a form of 'colonial legacy' bias. Similarly, Dreher & Fuchs (2015) documents that voting rights in international organizations can lead to biases in aid distribution.

Charitable giving has also been subject to studies about bias. DellaVigna et al. (2012) finds that social pressure could significantly affect charitable donations, and these biases could be effectively manipulated through planned marketing strategies. This evidence expands the scope of bias beyond individual and public decision-making to include non-profit sectors among others.

In a broader macroeconomic context, biases have been found to impact international capital markets. For example, Fuchs & Gehring (2017) finds systematic biases in sovereign rating decisions which influence a nation's borrowing costs and access to international capital markets.

Finally, biases are present in employment decisions too. The seminal work of Bertrand & Mullainathan (2004) opens the door to an increased understanding of racial bias in the labour market. By demonstrating that resumes with white-sounding names received more callbacks than those with African-American-sounding names, they discovered the existence of racial bias in employee hiring decisions. On the other side, Goldin & Rouse (2000) provides empirical evidence of gender bias, in the seemingly meritocratic setting of orchestra auditions. Their study revealed that the introduction of blind auditions led to a significant increase in the selection of female musicians, drawing attention to the role unconscious gender bias can play even in situations where decisions are believed to be made based purely on talent.

Although biases are found in several contexts in empirical studies, obtaining data from judges, reviewers, journal editors or other decision-maker decisions is quite challenging considering other factors also need to be controlled in analyses. In this, the sport may provide us with an opportunity to study referee decisions to see if there are any biases while deciding. This study examines referee behaviour by using data from football to test if social pressure could affect the decisions of individuals as claimed by Becker & Murphy (2000).

1.2.2 Bias in Sport

The sports literature on bias can be seen as a specific application of broader economic theories about how social forces, incentives, and information asymmetries can lead to biased outcomes. Many of the key theoretical concepts from the general literature on bias in economics have direct parallels in the study of sports.

One important theoretical connection is the idea that social pressure and social context can shape behaviour and decision-making. Akerlof (1980)'s concept of social distance and the identity economics framework of Akerlof & Kranton (2000) highlight how social context and identity can lead to biased outcomes. This has clear relevance in sports settings, where the social pressure from fans and the identity of being associated with a particular team have been shown to influence the decisions of referees and judges.

These identity-related biases are particularly evident when it comes to the nationalistic and racial biases exhibited by referees and judges in sports. Nationalistic bias, for example, was observed in football (B. R. Pope & Pope, 2015), ski jumping and figure skating (Krumer et al., 2022; Zitzewitz, 2006), where athletes received favourable treatment from referees and judges of their own nationality. Also, bias can originate from historical or political contexts, such as found in international football, where referees from post-communist states show favouritism towards teams from non-communist states, thus reflecting the influence of the Cold War (Dagaev et al., 2021).

Racial bias, on the other hand, operates subtly but systematically across different sports. This is evident in the evaluation of professional football players in Italian newspapers, where black players received lower ratings than their non-black counterparts (Principe & van Ours, 2022). A similar bias was observed in the NBA, with racial discrimination amongst NBA referees which have significant discrepancies in referee calls that favour players from their own race (Price & Wolfers, 2010). Furthermore, racial biases have been discovered in the sports labour market. For example, Hill & Remer (2020) finds that race can play a role in who will be hired, promoted, or kept in teams as a coach in the NBA. This means that the issue of racial bias is not only about the players and referees but also affects other stakeholders and almost all decision-making processes in different areas of sports.

Linguistic bias, another form of identity-related bias, has also been documented in sports. Faltings et al. (2023) finds that in Swiss football, referees issued significantly more yellow cards to teams that were not from the referee's linguistic area. They also found some evidence, in the highest-level league, of referees issuing more red cards to teams from different linguistic areas and away teams achieving fewer points when the home team shared the same linguistic area with the referee. The authors suggest that this bias is likely subconscious and reflexive rather than a deliberate act of discrimination, highlighting that identity-related biases can be very deeply rooted.

On the other side, there are studies showing that mitigating such biases is possible. For instance, in their study in international cricket, S. M. Chowdhury et al. (2024) highlights an unexpected change in the decisions of umpires. Historically, when umpires shared the same nationality as the home team, there was noticeable favouritism towards them. This bias led to the introduction of neutral country umpires. But with the Covid-19 pandemic, home umpires made a brief comeback while officiating behind closed doors. Thanks to the previously documented biases, higher scrutiny and the implementation of technology-driven decision reviews, these biases were mitigated. In fact, there appeared to be an overcompensation and close decisions more often went against the home team (S. M. Chowdhury et al., 2024). Such findings suggest that it is possible to tackle biases in sports through increased awareness and technological controls. However, such decisions could be seen as another type of bias as umpires tend to prove their impartiality while making biased decisions in favour of a team other than their own nationality. Furthermore, Dawson et al. (2020) discovers that the introduction of television match officials in rugby is associated with higher home bias exhibited by referees. This suggests that before the introduction of an additional official, who is remote from direct crowd influence, referees might have been either consciously or subconsciously avoiding home bias to prove their impartiality.

In the next part of this review, I narrow down the literature discussion to the referee bias in football, which is the main focus of the current study. The literature on referee bias in football can be categorised into two main areas: studies examining home bias and studies investigating favourite team bias. Home bias refers to the tendency of referees to favour the home team, while favourite team bias suggests that referees may be more lenient towards historically successful or popular clubs. This study aims to contribute to both areas of research by utilising a novel approach based on ball-in-play time.

Home Bias

Studies on referee behaviour in football can be grouped into two main categories. The first category, which is the primary focus of the current study, examines referees' decisions regarding additional time at the end of matches. Several studies have demonstrated that referees tend to favour home teams by adding more time when they are trailing, especially when the goal difference is only one at the end of the second half (Garicano et al., 2005; Sutter & Kocher, 2004; Dohmen, 2008; Scoppa, 2008; Rickman & Witt, 2008). These studies have not detected any bias in situations with a two-goal difference or more, as an additional goal during this time would not typically alter the outcome of the match. Therefore, it seems that referees tend to favour home teams only when their decisions could potentially affect the outcome of the match.

The second category of studies looks at referees' decisions regarding penalties, yellow cards, and red cards. These studies claim that referees tend to award more penalties to home teams and punish away teams more severely with yellow and red cards (Boyko et al., 2007; Dohmen, 2008; Sutter & Kocher, 2004; Dawson et al., 2007; Buraimo et al., 2010; Dawson & Dobson, 2010; Pettersson-Lidbom & Priks, 2010; Nevill et al., 2002). These findings suggest that referees make more favourable discretionary decisions for home teams across different aspects of the game.

The difference in referee decisions has been attributed to the pressure from home team fans in the stadium (Garicano et al., 2005; Nevill et al., 2002; Pettersson-Lidbom & Priks, 2010; Dohmen, 2008; Downward & Jones, 2007). Experimental evidence from Nevill et al. (2002) shows that referees make more decisions in favour of home teams when exposed to crowd noise. Moreover, Garicano et al. (2005) finds that increased stadium attendance leads to greater referee bias in favour of home teams, while Dohmen (2008) and Scoppa (2008) show that the presence of a running track in the stadium, which increases the distance between fans and the field, results in weaker referee bias. Lastly, studies fail to find evidence of such bias when matches are played behind closed doors due to the Covid-19 pandemic or other safety reasons, including stadium bans (Pettersson-Lidbom & Priks, 2010; Reade et al., 2022; Endrich & Gesche, 2020). These findings suggest that the social pressure created by the home team's fans in the stadium is a key factor influencing referee decisions.

Studies conducted during the Covid-19 pandemic have found a decrease in both the home advantage (Scoppa, 2021; Ferraresi & Gucciardi, 2023; Reade et al., 2022) and referee bias (Bryson, Dolton, et al., 2021; Scoppa, 2021; Reade et al., 2022) when matches were played without fans in stadiums. However, these studies often assume that the reduction in the home advantage is primarily due to the absence of referee bias, potentially neglecting the impact of the crowd on player performance.

However, it is important to note that the home advantage in football is a complex phenomenon that cannot be entirely attributed to referee bias. Factors such as travel fatigue and familiarity with the home stadium have also been cited as contributing to the home team's success (Courneya & Carron, 1992; Pollard, 1986; Pollard & Pollard, 2005). Ponzo & Scoppa (2018) finds that playing at home increases the probability of winning by approximately 25% although this advantage reduces to around 15% for teams competing against other teams using the same stadium. This suggests that factors beyond stadium familiarity, such as crowd support, may play a significant role in the home advantage.

Moreover, the influence of the home crowd may not be limited to referee decisions. Nevill et al. (1996) argues that the support of the home fans may also affect the performance and confidence of home team players, while the pressure from the crowd may lead to more aggressive play from away team players. Ferraresi & Gucciardi (2021) provides another evidence, finding that the home advantage is present even in isolated moments of the game, such as penalty kicks, where the missing probability of home teams is lower while that of away teams is higher when fans are present. Such instances require extreme concentration, and even the slightest influence from external factors like crowd noise or the perceived support or pressure of home fans can sway outcomes. Such factors may affect the playing styles of teams too. Therefore, away teams, to compensate for this disadvantage, may commit numerous and more violent fouls, which could cause yellow or red cards and potential penalties.

Similarly, Farnell (2023) provides evidence that the presence of fans can also influence player behaviour, finding that home teams in the NFL commit fewer defensive penalties, for which referees have no opportunity to show bias when fans are present. This highlights the need to consider the potential effects of crowd pressure on both referees and players when examining home advantage and referee bias. Therefore, neglecting any potential effects of fan pressure on players while examining referee bias in the number of cards, penalties, and fouls may lead to misleading conclusions.

Studies investigating referee bias may be neglecting the possible effects of the fans' pressure on players as they assume the toughness of actions is random or the same. Instead of analysing the number of fouls, cards and penalties awarded by referees, analysing such events one-by-one, and deciding if referees judge away teams more harshly or optimally, could disentangle other reasons for the differences in the numbers of fouls, cards and penalties between home and away teams. In fact, conducting such a detailed analysis would be challenging. However, when the toughness of actions is neglected and only focused on the number of them, it may cause a potential endogeneity problem stemming from omitted variable bias.

While the influence of home crowd pressure on referee decisions has been extensively studied, another dimension of referee impartiality that has received less attention in the literature is the potential bias towards traditionally successful and popular clubs. This 'favourite team bias' suggests that referees may be more lenient or favourable towards teams with a strong reputation or historical success. The following subsection explores the existing evidence on favourite team bias in football and other sports, highlighting the need for further research on this topic.

Favourite Team Bias

In addition to the home bias, a few studies have investigated whether referees exhibit favouritism towards teams that are perceived as stronger, more popular, or of higher status. Scoppa (2008) examines the teams involved in the 2006 Serie A scandal in Italy (Calciopoli) and found evidence of preferential treatment towards these clubs, with social pressure from the crowd identified as a primary driver of this bias. Similarly, Bose et al. (2022) analyses German football and showed that clubs with higher long-term status, as indicated by historical league rankings and membership numbers, benefited from fewer unfavourable decisions against them, even after controlling for their actual strength. However, Butler & Butler (2017) finds limited evidence of bias towards 'big' clubs in the English Premier League, suggesting that the extent of favourite team bias may vary across different football leagues and contexts.

The phenomenon of favourite team bias is not limited to football, as studies have documented similar patterns in other sports. For example, Mills (2014) finds evidence of umpire bias in favour of stronger teams in Major League Baseball (MLB), particularly during high-stakes games. Similarly, Zitzewitz (2006) shows that in winter sports, such as figure skating and ski jumping, judges tend to award higher scores to competitors from traditionally successful countries, potentially reflecting a bias based on the reputation and status of these nations in the sport.

These studies highlight that referee bias towards perceived stronger or higher-status teams is a widespread phenomenon that can manifest in various ways across different sports and leagues. The findings suggest that factors such as a team's historical success, popularity, and reputation may influence the decisions of referees, leading to more favourable treatment for these clubs. This bias can have important implications for the fairness and integrity of sporting competitions, as it may provide an undue advantage to certain teams based on factors unrelated to their current performance or abilities.

In summary, the literature on referee bias in football has identified two main forms of bias: home bias, where referees favour the home team due to the influence of crowd pressure, and favourite team bias, where referees exhibit preferential treatment towards traditionally successful and popular clubs. While the evidence for home bias is more extensive and robust, the research on favourite team bias is more limited and has found mixed results across different leagues.

The current study aims to contribute to this literature by providing new evidence on both home bias and favourite team bias in the context of additional time allocation in football. This study expands on Garicano et al. (2005) in several ways. It includes multiple European leagues for a broader perspective on referee bias while Garicano et al. (2005) uses data from only a single league, Spanish La Liga. The study also uses data from the Covid-19 pandemic as a natural experiment to isolate the effect of fan presence on referee decisions. While Garicano et al. (2005) finds about a 2-minute difference in additional time difference when home teams trail and win, this amount of bias could be biased as they were not able to control time wasting and other strategic stoppages. An important contribution of this study is the new methodological approach that uses ball-in-play data to calculate expected additional time. This method addresses the limitations of previous studies by accounting for all stoppages and time-wasting tactics, which simple event counts as in Garicano et al. (2005) may not fully capture. The findings of this study can help to inform efforts to promote fairness and impartiality in football officiating, as well as contribute to the broader understanding of how social and psychological factors can shape decision-making in sports and other domains.

1.3 Social Pressure and Referee Bias

In this section, I investigate the effect of social pressure on referee decisions. The Covid-19 pandemic provided a unique opportunity to test whether the presence of fans in stadiums influences referees' decision-making, as many football matches were played behind closed doors without fans. This natural experiment allows me to isolate the impact of social pressure on referee behaviour, specifically in the context of awarding additional time.

The study contributes to the literature by introducing a novel approach to quantify the effect of social pressure on referee decisions. I calculate the expected additional time based on the ball-in-play time in each half of the match, which serves as a benchmark for the amount of time that should be added to compensate for stoppages in play. By comparing the actual additional time awarded by the referee to this expected value, I can identify potential biases in referee decision-making and assess the extent to which these biases are influenced by the presence or absence of fans in the stadium. This approach mitigates potential omitted variable bias in the literature, as it focuses on referees' evaluations of required time, rather than any actions by the players.

1.3.1 Data and Methodology

Data for six European football[2] league game statistics, for the football seasons from 2018-2019 to 2022-2023[3], have been collected from the Football-Reference website.[4] Ball-in-play and additional time data have been provided by Stats Perform and Opta, which are industry-leading data providers.[5] I use actual additional time, which shows how many seconds a game is played, instead of the additional time indicated by fourth officials, as it is the minimum allowed time which referees can also extend if needed (IFAB, 2022).

The dependent variable under investigation in this analysis is the difference between the actual additional time awarded by the referee and the expected additional time. As previously outlined, the expected additional time is calculated for each half of the matches using ball-in-play. This expected additional time serves as a benchmark, established to account for deviations from the standard 62% ball-in-play time ratio. The intention behind using this benchmark is to ensure that the duration of a match aligns with generally accepted expectations, thereby negating instances where matches are either unfairly extended or shortened to favour a side.

This can be expressed mathematically as:

$$\Delta AT = AT_{actual} - AT_{expected} \tag{1.1}$$

where $AT_{expected}$ is calculated for each half using ball-in-play time:

$$AT_{expected} = (0.62 \times 2700) - BIP \tag{1.2}$$

Here, BIP represents the ball-in-play time in seconds, 2700 is the total number of seconds in a 45-minute half, and 0.62 is the standard ratio of ball-in-play time to total time.

[2]English Premier, German Bundesliga, Italian Serie A, Spanish La Liga, French League 1 and Turkish Super League.

[3]Until 18 September 2022.

[4]www.fbref.com

[5]The author thanks Stats Perform, Opta and Barış Gerçeker for providing ball-in-play data.

Previous studies analysing the decision-making process regarding additional time by referees have used the additional time added by the referee as the dependent variable, controlling for the number of stoppages such as substitutions, goal-kicks, and goals, among others, to account for time lost.While this approach offers valuable insights, it may not adequately estimate the referee bias. The duration allocated for each event can vary depending on the location within the match, the timing of these events, the current score or potential winner, teams' expectations, and the severity of the incidents (Morgulev & Galily, 2019). These factors, if not properly accounted for, could lead to biased estimates. For instance, VAR checks, goalkeeper injuries, or the location and timing of stoppages could significantly impact the additional time needed but may not be adequately captured by simple counts of events. These omitted factors could likely be correlated with additional time and score difference which is the variable of interest explained below. For example, as getting points in away games is harder and more important, leading away teams may engage in more time-wasting tactics which are not captured by simple event counts but affect the required additional time. Moreover, these time-wastings could be evaluated partially by referees depending on home or away teams.

In this study, I use the actual additional time that the game lasted, rather than the time indicated by the fourth official, which is used in previous studies. This approach addresses a potential source of measurement error, as referees have the discretion to extend the game beyond the indicated time if needed as stoppages during the additional time itself are also possible. Omitting controls like the side of the team committing fouls, taking goal-kicks or making substitutions, and the location of kicks (e.g., free kicks close to the goal may require more time to set up a wall) could bias the estimates and lead to misleading results. The approach in this study, using ball-in-play data to calculate expected additional time, addresses these issues of measurement error and potential omitted variable bias. By providing a more accurate measure of the time that should be added, this method reduces the likelihood of measurement error being correlated with the variables of interest, thus improving the consistency of estimates. By using ball-in-play data, this study inherently accounts for all these factors, capturing the actual time lost due to all stoppages, regardless of their nature or the tactics employed by teams. This comprehensive measure leads to more consistent estimates of referee bias and allows for a visual illustration of bias, as shown in Figure 1.2.

To ensure the robustness of the study, I conducted additional analyses using alternative ball-in-play ratios of 55%, 60%, 65%, and 70%. Furthermore, I replicated the analysis using the traditional approach from the literature, with actual additional time as the dependent variable and ball-in-play time as a control variable. These robustness checks presented in Tables A.1 and A.2 in the appendix support the reliability of the study.

By comparing the actual additional time awarded by the referee to this expected value, I can identify instances where the referee may have extended or shortened the match beyond what would be considered fair or appropriate based on the amount of time lost due to stoppages in play. A positive difference between the actual and expected additional time indicates that the referee has added more time than would be expected based on the ball-in-play time, while a negative difference suggests that the referee has added less time than expected.

It might be argued that each football match is unique, leading to variations in their duration. This variability could potentially impact the fairness of competitions. Differences in team playing styles contribute to this issue. Specifically, teams that adopt a strategy of deliberate slow play during throw-ins, goal-kicks, and other restarts might consistently experience shorter matches. This approach can advantageously reduce the physical strain on players, potentially resulting in fewer injuries. Therefore, to ensure fairness, it is necessary to enforce a standard playing time for every match.

The main independent variable is $Score\ Difference$ which is a dummy variable that takes 1 if the home team is leading and 0 if the away team is leading by 1 goal at the end of the 90^{th} minutes of the games as in Garicano et al. (2005). There are 9018 matches in the dataset, of which 3550 have a one-goal difference going into additional time.

Another important variable in this study is *Covid* which is a dummy variable, indicating whether a game was played during the Covid-19 pandemic. Table 1.3 presents the distribution of matches played during the Covid-19 pandemic by the leagues.

Table 1.3 *Games Played During the Covid-19 Pandemic*

		TOTAL					COVID-19				
	League	18-19	19-20	20-21	21-22	22-23	18-19	19-20	20-21	21-22	22-23
1	Bundesliga	306	306	306	306	63	0	83	305	0	0
2	EPL	380	380	380	380	67	0	92	363	0	0
3	La Liga	380	380	380	380	60	0	110	378	0	0
4	League 1	380	279	380	380	80	0	0	380	0	0
5	Serie A	380	380	380	380	70	0	124	380	0	0
6	Turkish	306	306	420	380	63	0	82	420	0	0

Note: League 1 was stopped in the 2019-2020 season when the pandemic emerged. Turkish League was played with 21 and 20 teams in the 2020-2021 and 2021-2022 seasons, respectively. Data source: Opta.

The key control variables in the model are the relative quality of the home and away teams, which is measured using pre-game betting odds data taken from Football-Bet-Data[6]. The rationale behind this choice is that betting odds are widely considered to be a reliable indicator of team quality, as they incorporate a wide range of relevant information, such as player injuries, suspensions, and recent form (Buraimo et al., 2010). Specifically, the odds ratio is calculated by dividing the odds for an away team win by the odds for a home team win. A higher odds ratio indicates that the home team is expected to be stronger than the away team, while a lower ratio suggests that the away team is favoured. Although betting odds are slightly adjusted according to the presence of fans after some time (Fischer & Haucap, 2022), they are still widely used in academic studies and seen as better control than alternatives including Elo ratings and Transfermarkt team values as betting odds better reflect the most precise strength of teams right before the games.

[6] www.football-bet-data.com

Furthermore, I investigate the effect of crowd size on the magnitude of potential referee bias by including attendance as a proxy for social pressure, following the approach of previous studies (Garicano et al., 2005; Pettersson-Lidbom & Priks, 2010; Buraimo et al., 2010; Downward & Jones, 2007). Garicano et al. (2005) and Nevill et al. (2002) suggest that referees may favour the home team due to the pressure exerted by the crowd, with larger crowds potentially leading to stronger bias. Downward & Jones (2007) finds that the probability of a yellow card being awarded against the home team decreases as crowd size increases. They propose two explanations: first, crowd noise may serve as a decision-making heuristic, with the likelihood of an incident being perceived as a foul increasing with crowd reactions (Nevill et al., 2002); second, referees may seek to appease the crowd, with this tendency becoming more pronounced as crowd size grows (Garicano et al., 2005). Table 1.4 presents the descriptive statistics of the variables used in the analysis.

Table 1.4 *Descriptive Statistics*

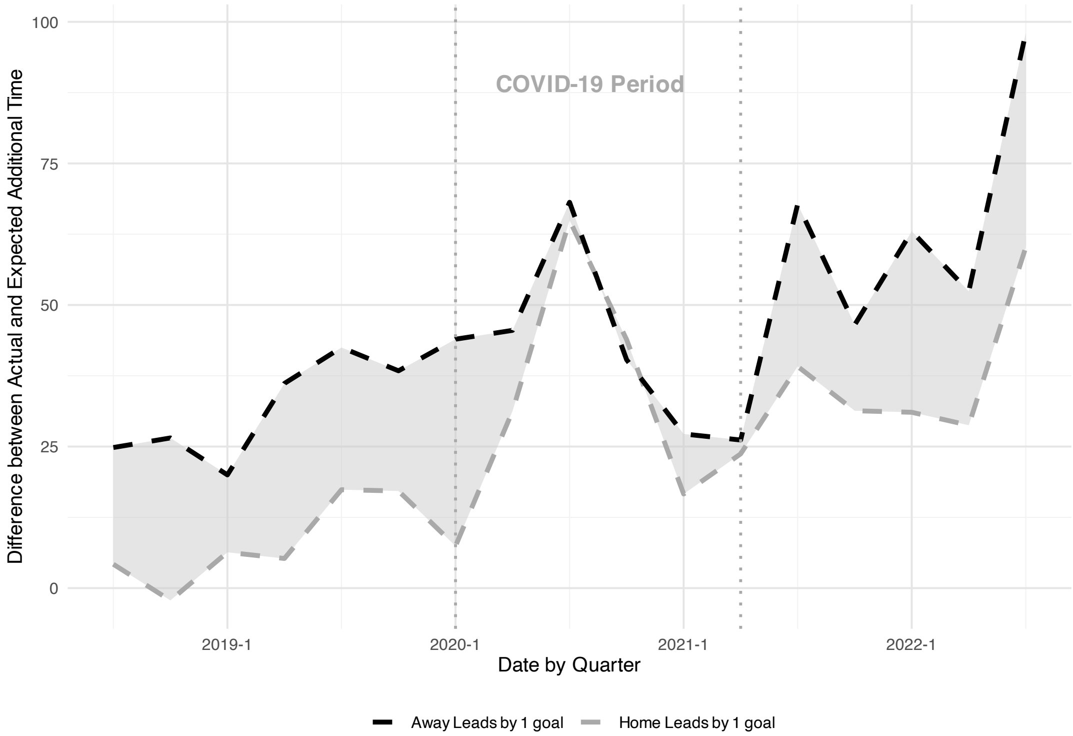
Variables	N	Mean/Share	St. Dev.	Min	Max
Difference in Score (Home−Away)	9,018	0.278	1.823	−9	9
Round	9,014	18.341	10.995	1	42
Attendance (number of people)	8,878	18,480	19,617	0	93,426
1st Half AT[a]	9,018	127.343	95.362	0	1,589
2nd Half AT[a]	9,018	275.314	111.671	0	1,065
1st Half Expected AT[a]	9,018	127.339	36.516	9	1,272
2nd Half Expected AT[a]	9,018	275.315	41.098	136	1,066
1st Half Ball-in-Play (in seconds)	9,018	1,668.905	168.000	1,019	2,226
2nd Half Ball-in-Play (in seconds)	9,018	1,630.471	165.729	1,040	2,221
1st Half Ball-in-Play Excluding AT[a]	9,018	1,619.122	187.041	873	2,224
2nd Half Ball-in-Play Excluding AT[a]	9,018	1,520.420	194.485	690	2,177
Quality Ratio (Away/Home betting odds)	9,018	2.556	3.401	0.043	44.485
Before Covid-19	9,018	0.407	0.491	0	1
During Covid-19	9,018	0.301	0.459	0	1
After Covid-19	9,018	0.292	0.454	0	1

Note: Data source: Opta. [a] Additional Time (in seconds).

The empirical strategy relies on a difference-in-differences (DiD) approach to estimate the effect of social pressure on referee decisions. The Covid-19 pandemic provides a natural experiment that allows to compare referee behaviour in matches played with and without fans in the stadium. By exploiting this exogenous variation in the presence of social pressure, I can isolate the causal effect of fans on referee decision-making. I define the treatment group as matches played without fans due to Covid-19 restrictions, and the control group as matches played with fans.

The key identifying assumption in this DiD framework is that the difference between expected and actual additional time in matches where the home team is leading or trailing by one goal at the end of the 90$^{\text{th}}$ minutes has followed parallel trends before the Covid-19 pandemic. This assumption is plausible given the sudden and unexpected nature of the pandemic, which led to the exogenous introduction of crowd restrictions. Also, Figure 1.2 illustrates the parallel trends before and even after the Covid-19 pandemic.

Fig. 1.2 *Covid-19, Presence of Fans and Referee Bias*



Note: Grey dashed line: Difference between the actual and expected additional time when home team is leading at the end of 90th minutes. Black dashed line: Difference between the actual and expected additional time when home team is losing at the end of 90th minutes. Shaded area between lines: Amount of bias. Vertical dotted lines and text indicate the Covid-19 period. Data source: Opta.

The DiD approach is implemented using the following regression equation:

$$\begin{aligned}\Delta AT_{ijkl} = \beta_0 + \beta_1 \times Score\ Difference_{ijkl} + \beta_2 \times Covid_{ijkl} \\ + \beta_3 \times (Score\ Difference_{ijkl} \times Covid_{ijkl}) \\ + \gamma \times \mathbf{X}_{ijkl} + \alpha_{Home \times Season_{il}} + \alpha_{Away \times Season_{jl}} \\ + \alpha_{Referee_k} + \alpha_{League_m} + \varepsilon_{ijkl}\end{aligned} \tag{1.3}$$

Where:

- ΔAT_{ijkl} represents the difference between actual and expected additional time for the game between home team i, away team j, officiated by referee k in season l.
- $Score\ Difference_{ijkl}$ is a dummy variable which takes a value of 1 if the home team leads the match by a one-goal difference and 0 if the home team trails by one-goal at the end of the 90^{th} minutes as in Garicano et al. (2005).
- $Covid_{ijkl}$ is a dummy variable indicating the presence of Covid-related restrictions on the presence of fans in stadiums.
- The interaction term, $Score\ Difference \times Covid$, is the main interest of the study and tests if the presence of fans affects referees' decisions by putting pressure on them.
- $\mathbf{X}_{ijkl}$ is a vector of game-specific control variables such as attendance, team quality ratio, if it is a weekday (see Krumer & Lechner, 2018), round (of the season) as later rounds become more important as compensation for a loss is limited.
- $\alpha_{Home \times Season, i(l)}$ and $\alpha_{Away \times Season, j(l)}$ are fixed effects capturing unobserved heterogeneity for each team in each season, $\alpha_{Referee,k}$ represents the fixed effect for the referee and $\alpha_{League,m}$ is a fixed effect for leagues to capture cultural difference of countries.

To test the relationship, I started by estimating a baseline OLS model that includes only the Score Difference variable and the fixed effects. I then gradually added the Covid dummy, the control variables and finally the interaction term,$Score\ Difference \times Covid$, to obtain DiD estimation. The coefficient of interest is β_3, which captures the effect of a one-goal lead on additional time allocation, comparing situations where the home team leads versus when the away team leads in matches played with versus without fans.

The main analysis of this study is a DiD estimation, instead of an OLS, thanks to the exogenous shock provided by the Covid-19 pandemic and the presence of parallel trends in the pre-pandemic period. Therefore, the estimates can be interpreted as the causal effect of the social pressure of fans on referee decisions regarding additional time allocation.

The DiD approach offers several key advantages over OLS regression in this study. Firstly, it allows for causal inference regarding the effect of fan presence on referee bias. The Covid-19 pandemic provides an exogenous shock to fan attendance, creating a natural experiment. DiD exploits this variation to isolate the causal effect of fan presence on referee decisions. OLS, in contrast, would only show correlation and could not distinguish whether fan presence causes referee bias or if other factors are driving both. Secondly, DiD effectively controls for time-invariant factors that might influence both fan presence and referee decisions, such as stadium characteristics, team reputations, or cultural factors specific to each league. While OLS with fixed effects could partially address this, DiD's before-and-after comparison provides a more robust control for these unobserved, time-invariant confounders. Thirdly, the DiD approach accounts for general time trends that might affect all matches, regardless of fan presence. For example, if there were a general trend towards more additional time being added in all matches over the years, DiD controls for this while OLS may mistakenly attribute such trends to the effect of fan presence. These advantages make DiD a more suitable method for estimating the effect of social pressure on referee decisions regarding additional time allocation.

1.3.2 Results and Discussion

Table 1.5 provides evidence of referee bias in favour of home teams. In the baseline model (column 1), I find that referees keep the game approximately 20 seconds shorter when home teams are leading by one goal at the end of the 90^{th} minute. This effect is statistically significant at the 1% level. The inclusion of control variables in model (2) does not substantially alter the magnitude or significance of the coefficient.

The interaction term in model (3) captures the effect of home fans on referee behaviour using the DiD approach. The negative and significant coefficient on $Score : Difference \times Covid$ indicates that the bias in favour of home teams stems from the presence of fans in the stadium. When fans are present, referees shorten the additional time by around 25 seconds when the home team is leading. However, in the absence of fans, there is no evidence of such bias. I test the hypothesis $\beta_1 + \beta_3 = 0$ and fail to reject it (p-value = 0.4778), confirming that referees do not exhibit any bias when fans are not in the stadium.

Moreover, the interaction term $Score\ Difference \times Attendance$ in model (4) shows a positive and significant relationship between the number of home fans in the stadium and the magnitude of referee bias. This finding aligns with the results of Garicano et al. (2005) and Downward & Jones (2007), suggesting that increased social pressure from larger audiences induces referees to favour the home team. The more fans are present in the stadium, the more likely referees are to shorten the additional time when the home team is winning. Considering that some big teams play in stadiums with a capacity of over 80,000, this effect could be substantial. For instance, if a home team with a stadium capacity of 80,000 is winning by one goal, the referee may shorten the additional time by around 63 seconds ($-0.7883 \times 80 = -63.064$) compared to a situation where the stadium is empty. This difference could be crucial in determining the outcome of close matches, as it gives the home team more time to equalise or even score a winning goal.

Table 1.5 *Effect of Home Fans on Referee Behaviour*

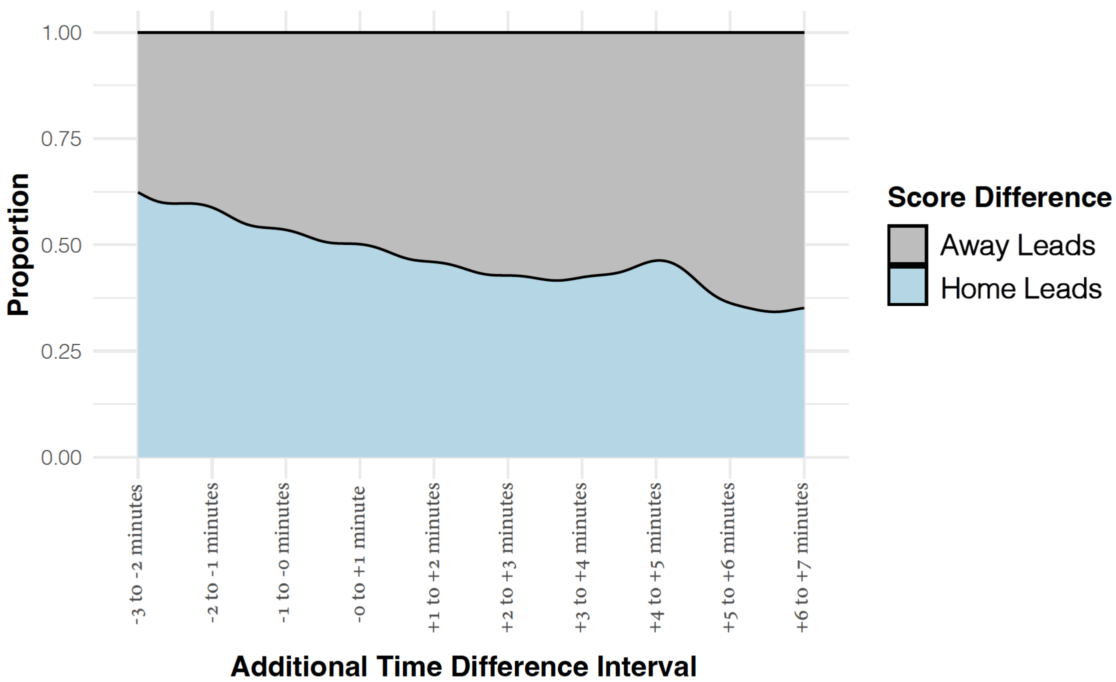
Dependent Variable:	*Actual − Expected Additional Time*			
Models:	(1)	(2)	(3)	(4)
Score Difference	-19.76***	-18.64***	-25.27***	-5.494
	(2.944)	(2.712)	(3.473)	(3.787)
Covid-19		27.44	17.53	29.11
		(19.313)	(20.036)	(20.093)
Attendance (1000)		-0.102	-0.058	0.431
		(0.296)	(0.282)	(0.337)
Weekday		2.697	2.640	2.501
		(3.272)	(3.234)	(3.285)
Round		-0.252	-0.234	-0.248
		(0.194)	(0.193)	(0.191)
Quality Ratio		-0.235	-0.262	-0.165
		(0.961)	(0.978)	(0.998)
Score Difference × Covid			20.43**	
			(6.386)	
Score Difference × Attendance (1000)				-0.788***
				(0.293)
Fixed-effects				
Home Team × Season	Yes	Yes	Yes	Yes
Away Team × Season	Yes	Yes	Yes	Yes
League	Yes	Yes	Yes	Yes
Referee	Yes	Yes	Yes	Yes
Observations	3,550	3,491	3,491	3,491
R^2	0.501	0.514	0.516	0.518

Data sources: Opta & Football-Reference. The dependent variable is the difference between actual and expected additional time(in seconds) at the end of the 90^{th} minute of matches that ended with a 1-goal difference as in Garicano et al. (2005). Four-way clustered (Home Team-Season & Away Team-Season & League & Referee) robust standard errors in parentheses. The decrease in observations from Model (1) to Models (2)-(4) is due to missing attendance data. * p<0.1; ** p<0.05; *** p<0.01.

The model applied individually to each league reveals a consistent trend in the top five European leagues; however, the Turkish Super League stands out, showing no significant referee bias in additional time decisions. Specifically, the magnitude of bias observed varies from 14 seconds in the German Bundesliga to 26 seconds in the Spanish La Liga, as shown in Table A.3 in the Appendix

Moreover, it is not just about adding 25 seconds more on average when home teams are trailing. The inclination of referees to extend the duration of matches goes beyond mere seconds; it often translates into several additional minutes. Figure 1.3 illustrates the probabilities associated with deviations in actual additional time from expected, depending upon which team, home or away, emerges as the winner at the end of the 90th minute, specifically in matches decided by a one-goal margin. For example, keeping games shorter by up to 3 minutes when home teams are ahead is possible. Table A.5 in Appendix shows the statistical results and magnitude of bias for each interval of additional time differences. Such a divergence from expected additional time could considerably influence the outcomes of matches, especially in tight games where every second counts. The presence of such a trend sheds light on the influence of home advantage and emphasises the importance of objective decision-making in the sport.

Fig. 1.3 *Probabilities of Deviations in Additional Time from Expected by the Winner by One-Goal at the End of the 90th Minutes*



Note: Probabilities of bias in intervals per minute, contingent upon the winner at the 90th minute. Data source: Opta.

In terms of outcome changes in the additional time exceeding the expected additional time, home teams manage to get more points by scoring one or more goals in 17.2% of the matches, while away teams only could benefit from the change in the outcome in 14% of the games. This statistically significant (p=0.0087) difference shows that home teams benefit from longer additional times even though relative team qualities which potentially increase the difference, are neglected.

Additionally, Table 1.6 separately estimates the referee bias before, during and after the Covid-19 pandemic to examine if referees develop immunity to external pressure. During Covid-19, when matches were held without fans, referees allocated additional time impartially; there was not any statistically significant difference between actual and expected additional time. Refereeing in the absence of fans and social pressure could be seen as a natural nudge environment. As biases are cognitive and unintended most of the time, referees can be expected to improve cognitive immunity against such social pressure during Covid-19 and continue making impartial decisions even when fans return to the stadiums. However, the additional time was approximately 31 and 24 seconds shorter than expected while home teams were leading by one goal, before and after the Covid-19 pandemic, respectively. Therefore, referees only marginally improved their impartiality after Covid-19, which shows the limited effect of such a natural nudge.

Although the NBA referees reduced their racial bias after being made aware of it through a study by Price & Wolfers (2010), as shown by D. G. Pope et al. (2018), football referees continued to exhibit home bias in the post-Covid-19 era. The discrepancy in the results could be attributed to NBA executives leveraging the study to enhance the impartiality of their referees' decisions, as racial bias is socially unacceptable and favours a specific group of players all the time. However, home bias does not carry the same sensitivity as racial bias. In round-robin tournaments, home bias does not favour any specific team or players, and both sides may experience both advantages and disadvantages in the long run though differences between relative team strengths could create impartiality which will be discussed below.

Table 1.6 *Referee Bias Before, During and After Covid-19*

Dependent Variable:	*Actual − Expected Additional Time*		
Models:	Pre-Covid-19	During Covid-19	Post-Covid-19
Score Difference	-30.74***	-2.128	-23.84***
	(4.905)	(7.312)	(8.642)
Round	-0.1293	-0.9371**	0.3016
	(0.2173)	(0.3071)	(0.3497)
Weekday	-5.196	9.102	-2.303
	(6.498)	(6.977)	(10.58)
Attendance (1000)	-0.0406	0.3088	0.1722
	(0.6188)	(3.378)	(0.3993)
Quality Ratio	1.316	0.4938	-1.965
	(1.687)	(2.326)	(2.338)
Fixed-effects			
Home Team × Season	Yes	Yes	Yes
Away Team × Season	Yes	Yes	Yes
League	Yes	Yes	Yes
Referee	Yes	Yes	Yes
Observations	1,405	1,082	996
R^2	0.562	0.637	0.651

Data sources: Opta & Football-Reference. The dependent variable is the difference between actual and expected additional time(in seconds) at the end of the 90th minute of matches that ended with a 1-goal difference as in Garicano et al., 2005. Four-way clustered (Home Team-Season & Away Team-Season & League & Referee) robust standard errors in parentheses. The variation in observation counts across models reflects the different time periods analysed: pre-Covid-19 (1,405 observations), during Covid-19 (1,082 observations), and post-Covid-19 (996 observations). * p<0.1; ** p<0.05; *** p<0.01.

Although studies suggest that referee experience (Nevill et al., 2002) and quality (Dawson & Dobson, 2010) inversely correlate with referee bias, this study fails to find a significant relationship between referee bias and factors such as referee age (as a proxy for experience) or quality (being a FIFA referee or officiating UEFA Champions League matches). The detailed results are presented in Table A.4 in the Appendix.

In summary, the findings of the study are in line with the studies of Endrich & Gesche (2020); Bryson, Dolton, et al. (2021); Reade et al. (2022) and Scoppa (2021). Studies examining the effect of social pressure on referee behaviour focus on the other decisions of referees, such as penalties, yellow/red cards and fouls. As discussed above, the absence of such social pressure may affect the behaviour of players (Farnell, 2023; Ferraresi & Gucciardi, 2021) besides that of referees and home players may play more violently to compensate for the absence of fans and home advantage stemming from the presence of their fans. Therefore, ball-in-play data provide robust evidence of the effect of social pressure on referee behaviour in additional time, while causality between social pressure and other referee decisions has limitations because of the potential endogeneity problems.

1.4 Team Strength and Referee Bias

In the second part, I examine if referees favour strong teams. In sports, perceptions of referee bias towards stronger teams have been a longstanding area of discussion. Team strength can be empirically captured using betting odds ratios and offer a quantifiable metric to represent a team's expected dominance in a match. By using this metric, this section seeks to empirically investigate the much-debated hypothesis: Do referees favour powerful teams? By examining this relationship, I aim to contribute to the favourite team bias discussions.

1.4.1 Data and Methodology

In this part, I employ the same dataset and a similar estimation strategy, with the dependent variable being the difference between the actual and expected additional time. I include $Score\ Difference$ for each score difference in the game along with an interaction with *Quality Ratio* to assess how the relative strength of a team influences the referee's decision on additional time, especially in close games. For top teams, even a draw against a weaker team may be perceived as a loss. By including scenarios where $Score\,Difference = 0$, it is possible to determine if referees tend to favour stronger teams in such matches by allocating more additional time, thereby providing them with more opportunities to score.

To examine the relationship between relative team strength, score difference, and the difference between actual and expected additional time granted by referees, the following regression equation is used:

$$
\begin{aligned}
\Delta AT_{ijkl} = \beta_0 &+ \sum_{d=-3}^{3} \beta_1 \times D_{d,ijkl} + \beta_2 \times Quality\ Ratio_{ijkl} \\
&+ \sum_{d=-3}^{3} \beta_3 \times (D_{d,ijkl} \times Quality\ Ratio_{ijkl}) \\
&+ \gamma \times \mathbf{X}_{ijkl} \\
&+ \alpha_{Home \times Season_{il}} + \alpha_{Away \times Season_{jl}} + \alpha_{Referee_k} \\
&+ \alpha_{League_m} + \varepsilon_{ijkl}
\end{aligned}
\tag{1.4}
$$

Where:

- ΔAT_{ijkl} denotes the difference between the actual and expected additional time for the game between home team i and away team j, officiated by referee k in season l.
- $D_{d,ijkl}$ is a dummy variable corresponding to each specific score difference, ranging from -3 to +3, for the given match.
- $Quality\ Ratio_{ijkl}$ signifies the relative team strength ratio for the specific match which is calculated using the method discussed in the first analysis.
- Interaction terms of $D \times Quality\ Ratio$ are the main interests of the study to estimate the effect of relative team strength for each score difference.

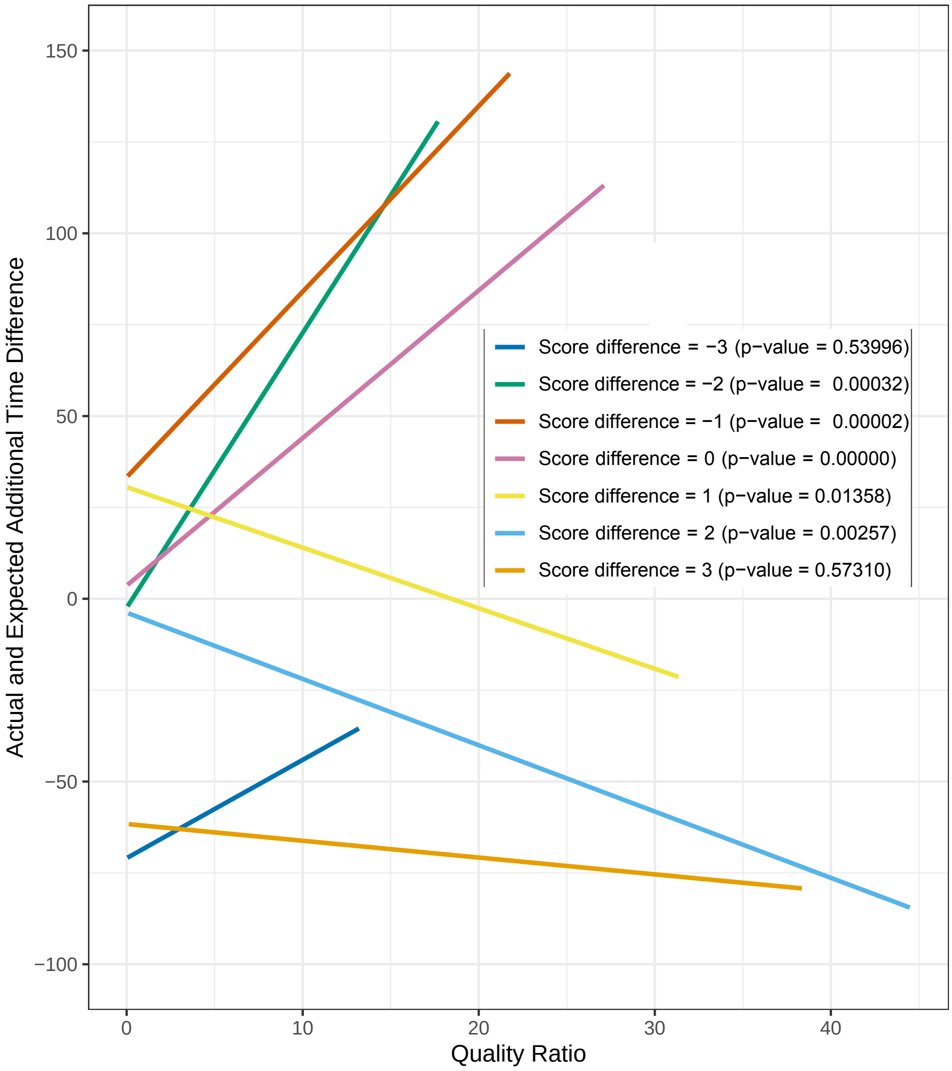
- $\mathbf{X}_{ijkl}$ is a vector of game-specific control variables, where each element represents a different control variable.

- $\alpha_{Home \times Season_{il}}, \alpha_{Away \times Season_{jl}}, \alpha_{League_m} and \alpha_{Referee_k}$ are fixed effects capturing unobserved heterogeneity specific to teams in a particular season, the league and the referee, respectively.

1.4.2 Results and Discussion

Figure 1.4 shows how the quality ratio relates to the difference between the actual and expected additional time for score differences ranging from -3 to 3. The figure illustrates that when the home team is stronger, referees tend to add more additional time than expected, especially if it provides the home team with an opportunity to tie the game or take the lead, typically seen in matches with a draw or a difference of one or two goals. Furthermore, when the stronger home team is already ahead, referees seem to shorten the additional time, possibly to decrease the chances of the home team conceding a late goal. However, these trends are not statistically significant for games with a goal difference of 3 or more.

Fig. 1.4 *Score Difference, Quality Ratio and Additional Time Difference*



Note: Data sources: Opta & Football-Reference. The figure displays multiple curves, with each curve corresponding to a different score difference. These score differences are detailed in the figure's legend with p-values. The higher the relative quality ratio between teams, the bigger the referee bias.

To further investigate the relationship between team strength, score difference, and referee bias in awarding additional time, I run regressions using the model specified above. The regression results, presented in Table 1.7, provide a quantitative examination of the results observed in Figure 1.4. The table includes two models: one for all games and another for games played exclusively during the Covid-19 pandemic. By comparing the results of these two models, the impact of fans' presence on the presence of referee bias towards stronger teams can be assessed.

Table 1.7 *Score Difference, Team Strength and Referee Bias*

Dependent Variable:	*Actual − Expected Additional Time*	
	All Games	Only Covid-19 Era Games
Model:	(1)	(2)
Covid-19	43.13***	
	(7.616)	
Attendance (1000)	0.2325	
	(0.1444)	
Round	-0.2915***	-0.2357*
	(0.0903)	(0.1369)
Weekday	6.255***	5.747
	(2.279)	(3.670)
Quality Ratio	0.3261	0.3784
	(0.6779)	(0.8711)
Score Difference (-3)	69.86***	60.34***
	(8.589)	(11.70)
Score Difference (-2)	148.3***	149.2***
	(7.577)	(9.481)
Score Difference (-1)	183.2***	183.0***
	(7.282)	(9.815)
Score Difference (0)	154.3***	155.8***
	(7.335)	(9.416)
Score Difference (+1)	169.7***	162.9***
	(7.714)	(10.67)
Score Difference (+2)	137.9***	137.7***
	(7.606)	(10.36)
Score Difference (+3)	75.06***	71.30***
	(7.596)	(10.32)
Quality Ratio × Score Difference (-3)	2.664*	0.0732
	(1.481)	(2.023)
Quality Ratio × Score Difference (-2)	3.978***	4.331***
	(0.9430)	(1.235)
Quality Ratio × Score Difference (-1)	3.484***	2.678**
	(0.9093)	(1.121)
Quality Ratio × Score Difference (0)	1.888**	1.428
	(0.8375)	(1.260)
Quality Ratio × Score Difference (+1)	-1.859***	-0.6677
	(0.6475)	(0.8772)
Quality Ratio × Score Difference (+2)	-2.016**	-2.084**
	(0.7809)	(1.046)
Quality Ratio × Score Difference (+3)	-0.3120	0.3777
	(0.7925)	(0.8692)
Fixed-effects		
Home Team × Season	Yes	Yes
Away Team × Season	Yes	Yes
League	Yes	Yes
Referee	Yes	Yes
Observations	8,874	3,668
R^2	0.473	0.466

Data sources: Opta & Football-Reference. The dependent variable is the difference between actual and expected additional time (in seconds) at the end of the 90th minute of matches. Four-way clustered (Home Team-Season & Away Team-Season & League & Referee) robust standard errors in parentheses. The variation in observation counts across models reflects the different time periods analysed: all games (1,405 observations) and only Covid-19 era games (3,668 observations). * p<0.1; ** p<0.05; *** p<0.01.

The results indicate that there is a referee bias in favour of stronger teams when determining additional time. Specifically, in matches where the potential for stronger teams to get more points exists, especially in matches with a goal difference of two or fewer, referees tend to allow longer additional time than expected. This bias is more pronounced when the strength disparity between the competing teams is wider. For example, when Italian giants Juventus were trailing Atalanta by one goal, additional time was 48 seconds more than expected. On the other side, while playing against Salernitana, additional time was 402 seconds longer and allowed Juventus to score a goal and get one point[7].

Furthermore, the analysis suggests that when stronger teams are leading, referees award shorter additional time than expected, arguably to secure the victory for the more dominant team. This deviation from the expected additional time depends on the relative strength of the teams competing in a match. For top-tier teams, a draw may have the same significance as a loss, given the high stakes of league standings and championship pursuits. This explains why referees award more time in such situations where stronger teams could potentially transition from a draw to a win, thus getting three points instead of only one.

When matches were played without fans during Covid-19, there is no evidence of the bias towards stronger teams, similar to the previous analysis. This suggests that the presence of fans plays a more significant role for stronger teams as the amount of bias gets bigger for stronger teams.

Some may argue that home bias is not a problem in round-robin tournaments, such as football leagues, because both teams experience it when playing away, and any advantages or disadvantages cancel out throughout the tournament. However, the presence of a bias towards stronger teams can interact with home bias and create an unfair environment. For example, if referees favour both home teams and stronger teams, then matches between a strong home team and a weak away team may be particularly imbalanced, with the

[7]Atalanta, a team that had competed in the UEFA Champions League that year, is considered superior to Salernitana by betting odds and by history.

stronger team benefiting from both biases. In contrast, matches between teams of similar strength may be less affected by referee bias. This interaction between home bias and team strength bias can lead to an unfair competition environment, where some teams systematically benefit more from referee bias than others. Buraimo et al. (2010) provides evidence for this interaction in their study of referee bias in the English Premier League and Bundesliga, showing that home bias is more pronounced when the home team is higher-ranked than the away team. Therefore, it is important to consider not just the existence of home bias, but also its potential interaction with other forms of bias when assessing the fairness of round-robin tournaments.

The findings of this study extend the existing literature on social pressure and decision-making in economics. The results demonstrate that the presence of home fans leads to biased decisions favouring the home team in football which aligns with the theoretical frameworks proposed by Akerlof (1980) and Akerlof & Kranton (2000) on how social context and identity influence economic choices. This effect is consistent with the work of Garicano et al. (2005) who found similar biases in Spanish football. However, this study advances this line of research by employing a novel measure of ball-in-play time, which provides a more precise estimation of referee bias.

Moreover, the analysis of the impact of team strength on referee decisions contributes significantly to the literature on favourite team bias, as explored by Bose et al. (2022). The observed tendency for referees to favour stronger teams, particularly in close matches, shows how strength can sway decision-making. The findings on team strength influencing referee decisions align with broader economic research on how status and reputation affect market outcomes as in sovereign ratings Fuchs & Gehring (2017). Covid-19 created a situation where matches were played without fans which allowed the study of decision-making without social pressure. The reduced bias observed during this period supports the ideas proposed by Becker & Murphy (2000) about how social environments affect individual behaviour. This reduction in bias is similar to findings in other areas. For example, Goldin & Rouse (2000) shows how anonymous processes in academia and blind auditions in orchestra hiring can reduce bias. These results highlight the importance of measurable ways to evaluate performance and make decisions in other industries. The use of expected additional time to identify referee bias in this study suggests that similar data-based methods could help reduce bias in other decision-making areas, such as in courts, when allocating resources, and in hiring, promotion and dismissal decisions.

1.5 Concluding Remarks

In this study, I analysed the effects of social pressure on human decisions. The study shows that football referees favour home teams and relatively stronger teams only when the pressure of home fans is present. The findings show that referees keep the game shorter while home teams are winning and longer while they are losing the game by one goal, but only in front of home fans. The magnitude of the effect depends on the relative team quality. The stronger the home team than the away one, the bigger the amount of bias. The bias holds as long as the goal difference is not so big, namely less than three. I conclude that social pressure has an effect on referee decisions as there is no such evidence when fans are absent. The study also shows that referees could not benefit from matches played behind closed doors to be able to continue deciding impartially after Covid-19. Promoting the issue in the media could raise the awareness of referees and help them to decide impartially as in D. G. Pope et al. (2018).

The study has a practical implication to prevent the unfairness which stems from referee bias. The governing bodies of football may utilise ball-in-play data to overcome referee bias in additional time. In the media, there have been ideas to improve the unfairness in additional time (e.g. Does Football Need a 60-Minute 'Stop-Clock'? (2022)). Based on these findings, this study proposes implementing an expected additional time system using ball-in-play data. This approach would involve automatically calculating additional time at the end of each half, based on actual playing time. Such a system offers several advantages: it eliminates subjective bias, provides transparency for all stakeholders, and addresses concerns about decreasing effective playing time. Importantly, this calculated additional time could be displayed on scoreboards throughout the match, allowing players, coaches, and fans to see it in real-time. While it may reduce some unpredictability in matches, the gains in fairness and consistency outweigh this concern. This method presents a balanced solution, retaining the traditional 90-minute format while ensuring fairer additional time allocation. Moreover, adopting this method could address the often complained concern of football coaches about decreasing playing time. It is recommended that football governing bodies consider this system to enhance the sport's integrity.

These findings about referee bias in football tell us something about decision-making in other areas too. The way referees react to crowd pressure might be similar to how other professionals, like judges or managers, respond to social influences in their work. This may imply that the findings from studying referee decisions in football can help us understand decision-making in other areas as well. For example, a judge could decide in favour of a local, wealthy or famous person in a case instead of being impartial. However, proving this would be challenging as such data is almost impossible to obtain. Therefore, using sports data, which is easy to measure and obtain, is beneficial for understanding general economic problems.

Limitations and Future Work

The study used additional time decisions of referees in football. Although the study disentangles other causes of referee decisions from referee bias and these findings could be generalised to other decisions of referees, evidence of bias from other decisions of referees including other sports might be helpful for the robustness of the study.

Although referee bias is documented in different areas and in different types, there is no study, to the best of my knowledge, investigating if teams or individuals who suffer from the bias, strategically respond to such biased decisions. Future research might explore whether disadvantaged groups can strategically counteract such biases, and conversely, if advantaged groups intentionally act to amplify and exploit these biases.

Furthermore, examining the impact of technological interventions, such as the introduction of Video Assistant Referees (VAR), on referee bias is a critical area for future study. The VAR system allows for the review of important decisions, including penalties and red cards, with the potential to enhance decision-making accuracy. Investigating whether VAR effectively reduces referee bias, or if biases continue to influence decisions despite this technological aid, could provide crucial insights into the role of technology in ensuring fairness in refereeing.

Chapter 2

Captains vs. All-Stars: Who Makes Better Leaders?

2.1 Introduction

Are workers paid according to their marginal productivity? The classic economic model claims that workers are paid according to their marginal productivity of labour. However, in team environments, this becomes more complex as workers contribute not only directly, but also indirectly by influencing the productivity of their teammates and colleagues (Arcidiacono et al., 2017). Also, some workers may be more talented in terms of influencing and leading others using their knowledge or charisma (Hermalin, 1998). However, it is not always easy to understand the contribution of individuals to team production in many industries because of the scarcity of data. Team sports can offer invaluable insights about the effect of peers on individual and team productivity to understand such dynamics as sports competitions provide team and individual-level time-varying observable productivity data.

Peer effects have been analysed in different areas ranging from academia to workplaces. However, studies find contradictory results. While some find a positive effect, others find negative or no significant effects. Such differences have been explained by the different mechanisms in the contexts of the studies.

Data from sports competitions have been extensively used by researchers to understand peer effects (e.g., Guryan et al. (2009); Arcidiacono et al. (2017); Jane (2015)). A comprehensive review of these studies will be provided in the literature review section.

On the other side, hiring superstars and selecting team captains are important decisions for sports teams as most of the time such kinds of players are seen as leaders who can increase the performance of their teammates. Superstars, known for their exceptional skills and marketability, are usually paid high salaries, raising questions about their effect on their teammates. It is believed that thanks to their talent, knowledge and charisma, they can be a source of spillover or convince others to perform better. Moreover, team captains are expected to lead and motivate to enhance the team performance. However, the efficacy of team captains and their impact on team success have not been empirically explored, yet.

The effect of superstars in sports has been analysed empirically on various aspects. Studies have examined the impact of superstars on peer performance, showing that their presence can motivate (Bilen & Matros, 2023; Brown, 2011) or demotivate (Deutscher et al., 2023) competitors, and have positive spillover effects on teammates (Wegelin et al., 2022). Superstars also influence economic factors such as stadium attendance (Bryson et al., 2014; Humphreys & Johnson, 2020; Jane, 2016; Schreyer & Singleton, 2023; Chmait et al., 2020; Kaplan, 2022), attendance at away games (Berri & Schmidt (2006)), and TV audience ratings (Wills et al., 2022; Reilly et al., 2023). However, the empirical research on the contribution of team captains to their teams is limited. To date, only sports psychologists have studied the characteristics and importance of team captains using qualitative research techniques (Fransen et al., 2014; Butalia et al., 2021; Santos et al., 2019).

In this paper, I investigate the peer effects of formally assigned leaders (team captains) and informal leaders (All-Stars) by using in-game injuries as random exogenous shocks (Wegelin et al., 2022) in staggered difference-in-differences estimation (Callaway &

Sant'Anna, 2021). When discussing leadership in sports teams, many people often think of team managers or head coaches. However, there can also be leaders among athletes and players themselves (Fransen et al., 2014). Although athlete leadership is a very popular topic among sports psychologists, empirical research on the effects of athlete leaders on their teammates is scarce. This gap is particularly pronounced in understanding formal leadership roles, such as team captaincy, interact with the informal influence of talented players, who could be All-Star players. The study aims to bridge this gap by offering new insights into how different forms of leadership roles in high-stakes team environments. By utilising team captains as formally assigned leaders and All-Star players as informal leaders, I will analyse their impact on team and individual performance in the NBA.

Basketball is a team sport played between two teams of five players each. The objective is to score points by shooting the ball through the opponent's basket. A successful shot from inside the three-point line, which forms an arc around the basket, scores two points. Shots made from beyond this line are worth three points. When a player is fouled in the act of shooting or if the opposing team exceeds the foul limit, the fouled player is awarded free throws. Each successful free throw is worth one point. The cumulative points from these various types of shots determine a team's score. A game in the NBA consists of four 12-minute quarters. In basketball, a 'play' refers to a single offensive possession. It begins when a team gains control of the ball and ends when they score, lose possession, or the game clock expires. In this study, I use play-by-play data which provides a detailed, chronological record of each play during a game. This includes information such as which players are on the court, who attempts shots, scores points, commits fouls or makes assists. Each play typically lasts between 10 to 24 seconds, depending on the pace of the game. These detailed data allow for a precise analysis of player performance and team dynamics throughout a game including how these may change in response to events like player injuries or substitutions.

This study combines a novel methodological approach with in-game injuries as unexpected and exogenous shocks as in Wegelin et al. (2022) and fouled outs[1], providing a natural experiment setting to observe the impact of superstars and team captains on other players in real-time using play-by-play game data. This method allows me to isolate the immediate effects of such players' absence. This study will help us understand if hiring a superstar affects the performance of incumbents in a team of workers. On the other side, are team captains able to help their team members to perform better? If this is possible, are teams able to choose their team captains effectively?

The implications of the findings extend beyond the scope of professional sports, as they may offer insights for team management across different industries. Understanding the effects of leadership can inform strategies for team composition, leader selection, and performance optimisation. For sports teams, this could translate into more informed decisions regarding captain selections and talent acquisitions. In other corporate environments, these insights could guide leadership development and team-building strategies.

The remainder of this chapter is organized as follows: Section two presents an extensive review of the literature by benefiting from studies on peer effects, superstars, and leadership across both economics and sports. Section three differentiates between captains, All-Stars, and other players in terms of their performance. Section four examines the impact of leadership players' absence on the performance of their teammates and team productivity. The paper concludes with section five, which summarises the key findings and implications of the study for sports teams and general team environments.

2.2 Literature Review

The literature review is organised in three key sections: Peer Effects, Superstars, and Leadership. In each section, I explore research findings from two fields: economics and

[1] In the NBA, a player is disqualified from the game after accumulating six personal fouls.

sports. This approach allows us to understand the findings of scholars in both disciplines about these topics.

2.2.1 Peer Effects

In this section, I will review the literature on peer effects in both economic and sports literature. This concept examines how the behaviour, performance, and decisions of one person can affect other people around them. By comparing studies from economics, where peer effects might involve how workers' productivity influences their colleagues, to sports, where a team member's attitude or performance can impact the other players in the team, I aim to understand similarities and differences by getting insights from both fields. This comparison will help us to understand the power of peer effects from different contexts and how it shapes team dynamics.

Peer Effects in Economics

In economics, the study of peer effects has shown the complicated ways in how individuals influence each other's productivity and behaviour across diverse settings. For example, Mas & Moretti (2009) uses field data from a large supermarket chain and exploits the natural variation in the composition of workers across shifts to study peer effects among cashiers. They find evidence of positive spillovers from high-productivity workers to their peers. Bandiera et al. (2010) analyses field data from a fruit-picking farm and exploits the natural variation in the allocation of workers to different rows to investigate peer effects in a low-skilled work environment. They find that workers are more productive when working alongside more able colleagues.

Azoulay et al. (2010) contributes to the conversation by studying the 'superstar extinction' phenomenon, where the untimely death of leading scientists diminishes their collaborators' output. Similarly, Herrmann & Rockoff (2012) examines the impact of teacher absences on student performance, revealing significant productivity losses when regular teachers are replaced. Moreover, they find that the effect is greater for more experienced teachers and short-term replacements as school managers were able to find experienced teachers to substitute for long-term absences. Conversely, Waldinger (2012) employs an instrumental variables approach to study spillover effects resulting from the displacement of scientists during the Nazi regime in Germany and finds no evidence of peer effect.

In their controlled experiment, Falk & Ichino (2006) reveals that peers can substantially effect productivity, with the effect varied by the competitive or collaborative environment in teams. Furthermore, Brady et al. (2017), in their natural experiment, explores the negative side of peer effects on student achievements in the US Naval Academy by showing how low-performing individuals can detrimentally affect group performance which highlights the potential negative consequences of peer effect.

Ultimately, Herbst & Mas (2015) examines both experimental and field studies to assess the impact of peer effects. Their analysis reveals that, although a significant number of studies find positive effects, there are also studies which report negative or no significant peer effects. This suggests that the peer effects are complex and can vary depending on the specific context or conditions of the study.

As a result, these studies demonstrate that peer effects can have significant impacts on productivity and performance across various settings, although the direction and magnitude of the effect may vary depending on the specific context. From enhancing productivity in retail and agriculture to driving innovation in academic studies, peer effects are an important aspect of human interaction that shapes economic activities. Understanding such dynamics offers valuable insights for organisational management.

Sports provide us with an opportunity to estimate the peer effects in team settings with their easily measurable and available performance data. One can argue that evidence from sports is less relevant because it represents only a specific part of our lives. However, evidence from sports is as valuable as evidence from fruit-picking (Bandiera et al., 2010) or cashiers (Mas & Moretti, 2009) which are also quite specific to certain contexts and relevant to only a few individuals.

Peer Effects in Sport

The literature on peer effects extends into sports as well, where the effect of an individual on others offers valuable insights into the understanding of peer effects thanks to the availability of observational data and the high-stake context of professional sports. This literature review looks at several sports disciplines, examining how peer effects exist in each sport and contribute to our understanding of it.

Basketball, with its rich team dynamics, serves as a fertile ground for analysing peer effects. Wegelin et al. (2022) and Arcidiacono et al. (2017) use play-by-play data from the NBA and emphasise the positive spillovers from high-performing players to their teammates. Pazzona (2022) furthers these findings by showing how a player's productivity could be increased in the NBA after playing with high-skilled peers in the Olympics. On the other side, Lackner (2023) investigates how the presence of a dominant competitor affects the decisions of other competitors and finds positive effects when dominance is reduced and negative effects when there is a clear dominant team on the effort of other teams in national team competitions such as Olympics and FIBA World Cup.

Studies by Guryan et al. (2009) and Brown (2011) examine dynamics of peer effects in golf. Guryan et al. (2009)'s investigation reveals that the performance of playing partners does not significantly affect a player's performance which suggests that peer effects are limited in individualistic sports. Conversely, Brown (2011) highlights how the withdrawal of a superstar competitor can motivate remaining players to perform better for the chance of winning the tournament in tennis.

Gould & Winter (2009) investigates the peer effects in baseball and finds that a player's batting average tends to go up when his teammates are also batting well, but it drops when the team's pitching quality is high. On the other hand, a pitcher's performance improves with better pitching from teammates.

Track and field, characterised by both individual and team events, offers a distinct perspective on peer effects. Depken & Haglund (2011) investigates 4×400 men's relay teams and finds that higher-quality runners underperform relative to their expected performance when they use individual performance as a benchmark. Moreover, Emerson & Hill (2018) explores if the presence of pace setters affects athletes in marathon races and finds negative effects.

Swimming, another sport combining individual and team efforts, provides additional insights. Jane (2015) demonstrates that the presence of high-performing peers of student-athletes in training and competitive settings can lead to significant performance enhancements among swimmers. Jiang (2020) confirms this finding at the professional level but only for female athletes. Research by Yamane & Hayashi (2015) finds that athletes swim faster when their peers swim behind them than when swimming alone and slower while swimming behind their peers. They claim that the key determinant of peer effects is observability. Social pressure is explained as the potential reason for higher effort and peer effects in these contexts. Neugart & Richiardi (2013) extends this understanding by examining relay teams and shows that earlier swimmers in teams free-ride on contribution to team production and swim slower than their individual performance while later swimmers perform better.

In football, studies have observed positive peer effects. Cohen-Zada et al. (2023) finds that players' effort substantially enhances their teammates' effort level and team production using distances covered by players. Ichniowski & Preston (2014) shows that the performance of a player improves after joining a better team. Similarly, Molodchik et al. (2021) demonstrates that football players exhibit enhanced performance and receive higher FIFA video game ratings when part of a stronger team.

Lastly, Mao (2023) shows that players tend to shirk in the presence of teammates with similar roles in esports[2]. Hoey et al. (2023) finds a decrease in both outputs per player and total team productivity when a complimentary player is absent because of an injury in hockey.

To sum up, these studies across diverse sports disciplines increase our understanding of peer effects by showing how the presence and performance of peers can significantly affect individual and team productivity. The wide range of studies highlights that peer effects are important in many different sports, but they are in different ways depending on the context. This knowledge could be beneficial for athletes, coaches, team managers, and sports psychologists working towards better team and individual athlete performance. By understanding the specific mechanisms through which peer effects operate in their respective sports, stakeholders can develop targeted strategies to foster positive peer interactions and mitigate potential negative influences. This may involve optimizing team composition, training practices, and leadership roles to create an environment that encourages constructive peer effects and enhances overall performance.

2.2.2 Superstars

In this part, I will briefly review the studies on superstars in economics and sports to understand their importance for teams.

Superstars in Economics

To explore the concept of superstars in economics and management research, few studies offer valuable insights into the effects of superstars across different industries. First of all, Rosen (1981) introduced the idea that minor differences in talent could result in significant differences in earnings and recognition, establishing a basis for the study of the superstar phenomenon in the field of economics. This concept was further criticised by Adler (1985),

[2] Dota 2

who defined a superstar may also be an individual whose talents are not only superior but also a popular public figure who is rewarded for that.

In businesses, Malmendier & Tate (2009) focuses on the consequences of CEOs attaining superstar status on the performance of their companies. By using prestigious business awards as a proxy for CEO status, they find that CEOs underperform in their roles after receiving an award compared to their earlier achievements and compared to similar CEOs who did not receive awards. Despite this lower performance, these superstar CEOs frequently receive higher pay afterwards and become more involved in external activities, such as joining other boards and writing books. Additionally, they identify an increase in earnings after the receipt of awards, especially in companies with weaker governance structures which suggests that the media-induced superstar status may not always align with shareholder interests.

Moreover, Azoulay et al. (2010) examines the impact of star scientists on their peers' productivity and shows significant decreases in output after the unexpected death of superstar researchers. This finding highlights the important role of superstars in fostering high productivity and innovation.

These studies show the role of superstars in economics and management by emphasising the detailed effects of exceptional individuals on performance, compensation, and organizational dynamics in different industries.

Sport also presents superstars who earn possibly more than CEOs and superstar academics do. Also, their capacity to be an idol for the masses and their colleagues is greater. Therefore, superstars of sports are matter as superstars of other areas including CEOs (Malmendier & Tate, 2009) and scientists (Azoulay et al., 2010).

Superstars in Sports

In sports, the concept of 'superstars' plays an important role in shaping team dynamics, influencing fan engagement, and driving economic outcomes. This part reviews the findings from several studies that examine the impact of superstars in sports which ranges from peer effects within teams to economic impacts such as stadium attendance, away game attendance, and television audience ratings.

Deutscher et al. (2023) focuses on the shadow effects of superstars and reveals that the presence of a superstar in tennis tournaments can demotivate competitors even in the earlier rounds before competing with them. Similarly, Brown (2011) and Bilen & Matros (2023)'s studies extend this discussion to golf and chess, respectively and show that superstars not only affect their immediate peers but also shape the competitive landscape and outcomes of tournaments with their existence.

The effect of superstars goes beyond the team and competition dynamics by significantly affecting fan engagement and other economic activities. Studies on stadium attendance, such as Bryson et al. (2014); Humphreys & Johnson (2020); Jane (2016); Schreyer & Singleton (2023); Chmait et al. (2020), consistently demonstrate that superstars attract larger crowds, both at home and away games in sports ranging from football to basketball and tennis. Moreover, Brandes et al. (2008) distinguishes between the fan attraction capabilities of national superstars and 'local heroes', star players on teams without national superstars, using nine years of attendance data from the German Bundesliga and find that national superstars boost attendance at both home and away games while local heroes primarily increase attendance at home games only.

The superstar effect is also effective in the domain of television audience ratings, where the presence of such players can significantly increase the number of audiences. Wills et al. (2022) and Reilly et al. (2023) show how superstars not only enhance the viewing experience but also increase the commercial value of broadcasts.

As a result, superstars emerge as main figures who not only affect the performance of their peers but also serve as key drivers of fan engagement, market demand, and financial success in the sports industry. These studies emphasise the importance of superstars in shaping the competitive structure of sports by contributing to both performance and economic indicators.

2.2.3 Leadership

Leadership in Economics

Studies on contract design and incentives in economics have provided an understanding of the conditions under which transactional leadership succeeds or fails, as explained by works such as Jensen & Meckling (1976); Hart & Holmström (1987) and Holmstrom & Milgrom (1991). These studies show the importance of factors in determining the efficacy of contractual agreements in leadership dynamics in workplaces. However, this focus has often neglected the potential of non-transactional (transformational) in influencing follower behaviour, a gap in the literature that scholars have begun to address only recently.

Benjamin Hermalin's studies (Hermalin, 1998, 2013) mark a significant shift in the understanding of leadership from purely transactional to transformational ones. His initial focus on the importance of trust and communication between leaders and followers (Hermalin, 1998) evolves into a broader perspective by emphasising the leader's role as a model for behaviour to boost morale and affecting the emotional states of followers to create a corporate culture (Hermalin, 2013). This progression highlights a move towards leadership which blends transactional and transformational leadership skills.

Transactional leadership emphasises the achievement of tasks and goals through mechanisms of rewards and penalties. For example, in a political context, a leader might employ transactional tactics to ensure party members' support for key policies by offering political favours in return for loyalty and success (Burns, 1978). Similarly, in the educational context, a school principal might set clear performance targets for teachers and may reward those who meet these objectives to enhance school performance (Pearce et al., 2003). This approach could be effective in achieving immediate results but may overlook the needs of individuals within the organization.

On the other side, transformational leadership seeks to inspire and motivate followers to achieve higher levels of performance by fostering a shared vision and personal development (Bass, 1985). In management, a transformational CEO might focus on cultivating a culture of innovation and collaboration by encouraging employees to take initiative and contribute to the organization's vision (Bass et al., 2003). For example, in academia, transactional and transformational leadership can be observed through the roles of a department head and a superstar researcher. The department head as a formal leader may use transactional leadership techniques and focus on setting clear goals by managing resources and rewarding or punishing faculty members using grants, promotions and allocating workloads. In contrast, a superstar researcher may informally emerge as a transformational leader by inspiring colleagues and students thanks to their knowledge and experience and may shape their research and teaching. Transformational leadership not only enhances performance but also contributes to the personal development and job satisfaction of employees without reward and punishment mechanisms.

Building on the principles of transformational leadership, the concept of identity economics, as introduced by Akerlof & Kranton (2000, 2005, 2010), extends the idea of leadership beyond traditional measures of performance to include the shaping of followers' identities and preferences. This approach aligns closely with the transformational leadership style, which emphasises inspiration, motivation, and personal development. In their studies, they use the military as an example of being a part of an identity and explain that intrinsic motivation derived from the sense of identity and belonging could be more valuable than extrinsic rewards to achieve goals. This shows the significant impact of transformational leadership not only in achieving immediate goals but also in building a cohesive and motivated team which is quite important for sports teams.

Leadership in Sport

In sports leadership, there is a complex situation regarding the roles and effects of different leadership figures in teams, including coaches, team captains, and informal leaders. This literature review briefly synthesises key findings of athlete leadership studies that focus on the implications for team performance, athlete motivation, and the psychological aspects of leadership in sports contexts.

Fransen et al. (2014) challenges the belief that team captains are always the most effective leaders. They claim that the optimal leadership structure within sports teams may not always align with traditional roles. This is supported by their survey findings from nine different sports, which show that almost half of the participants do not perceive the team captain as a leader. Butalia et al. (2021) also provides similar evidence, questioning the efficacy of current methods for selecting team captains. They discuss a potential mismatch between required leadership skills and the roles assigned in football and volleyball teams. These findings require a re-evaluation of how leaders must be chosen in sports teams and call for a more strategic approach that aligns leadership capabilities and team needs.

Cotterill & Fransen (2016) focuses on the distinction between formal and informal athlete leadership and emphasises the complementary roles they play in shaping team dynamics and increasing performance. The study highlights the importance of recognising and nurturing informal leaders within teams, who often exert significant influence through their actions, expertise, and interpersonal relationships. Fransen et al. (2015) further emphasises this point by suggesting that the influence of athlete leaders can exceed that of coaches especially when fostering team cohesion and a shared sense of purpose and identity in sport teams.

Moreover, the implications of athlete leadership for the performance and confidence of players in teams were analysed by sport psychologists through surveys. For example, Fransen et al. (2016) and Santos et al. (2019) examine how leaders' confidence and their ability to convey a sense of team purpose contribute to enhanced performance outcomes and they find that effective leadership raises confidence, performance and the sense of identity of athletes in sports teams.

According to these findings, teams can create a team environment by choosing the right leaders in teams and promoting an identity which helps them to gain confidence and be more successful in their competitions.

This study extends previous research on peer effects and leadership in sports, particularly building on the works of Arcidiacono et al. (2017) and Wegelin et al. (2022). Arcidiacono et al. (2017) uses play-by-play data between the 2005-2006 and 2008-2009 NBA seasons to examines productivity spillovers among teammates. They employ a production function approach and find that players' productivity increases when playing alongside high-performing teammates. However, they find that player compensations are primarily based on direct contributions to the team production and only little weight is given to their ability to increase their teammates' productivity. Wegelin et al. (2022) uses in-game injuries as exogenous shocks to study the impact of high-performing players on team productivity. They use play-by-play data from 15,707 NBA games between the 2004-2005 and 2016-2017 seasons and employ a canonical difference-in-differences method which assumes that the effect of every injury is the same neglecting its time. Their findings indicate that the absence of high-performing players decreases the field goal percentage of their teammates by 1.08%. However, this finding has limitations. As they are not able to control the quality of opponent players, they assume that the strength of the defence is the same whether a high-performing player is in the game or not. However, the opponent team may adjust their strategies and can employ tougher and softer defence strategies after a high-performing player gets injured.

In this study, I use play-by-play data from NBA as in Arcidiacono et al. (2017) and Wegelin et al. (2022) and employ staggered difference-in-differences as the effect of injuries could be different depending on their timings (Callaway & Sant'Anna, 2021). Moreover, I control for the quality of opponents by using ESPN's real plus-minus which is a limitation of previous studies. On the other side, while earlier studies examined the impact of high-performing players on team productivity, this study specifically focuses on the role of formally assigned leaders (team captains) and informal leaders (All-Stars) in basketball.

By distinguishing between these different types of leaders, this study provides insights into the relative importance of formal authority versus recognised talent in influencing player and team performance. This approach not only advances our understanding of peer effects in sports but also contributes to the literature on leadership and team dynamics in high-pressure environments. The findings of this study have implications for team composition and leadership selection strategies not only in sports but also in other collaborative and high-stakes environments.

2.3 Leaders vs. Others

In this part, I begin with a preliminary analysis to see if leader players (captain and All-Star) are better performers than non-leaders as there could be a potential spillover of performance which can enhance the performance of non-leaders. Team captains are formal leaders of teams assigned by coaches or team managers. On the other side, All-Star players are exceptionally talented and played in the All-Star game. The All-Star game is a single game that is a showcase of talent in the NBA as the most talented players are chosen by experts and public votes every year. Some players are chosen thanks to their exceptional talent while others thanks to their popularity in the media and public. In that manner, we can consider All-Star players as superstars according to definitions of both Rosen (1981) and Adler (1985). The next part compares the performance of player types before the examination of whether such players can affect performance of their teammates.

2.3.1 Data and Methodology

For this part, I use player-match level data from the NBA seasons from 2002 to 2021 to see the overall individual contributions of players to team productivity. I obtained the data from ESPN which has been the broadcaster of the NBA for more than the last two decades. I use ELO ratings of teams to control for team qualities, rather than the betting odds used in Chapter 1. This is because basketball betting odds present challenges. In basketball, bookmakers often use point spreads, which represent the expected margin of victory and are designed to create equal betting on both teams. This practice means that the odds are heavily adjusted based on team qualities which potentially makes them less reflective of true team strength compared to other sports. Standardised betting odds for basketball are only consistently available from 2008 onwards. For robustness, analyses using Elo differences are replicated using betting odds ratios (opponent odds divided by team odds), obtained from OddsPortal, as a control for team qualities. These robustness checks, presented in tables B.1 and B.2 in the appendix, confirm the findings.

Summary statistics of match-player data are provided in Table 2.1. To test if leader players perform better than others, I use the OLS regression equation given below. The main dependent variable is Real Plus-Minus (RPM) while other boxscore metrics are also used to understand in-game dynamics.

$$\begin{aligned} \text{Performance}_{imst} = {} & \beta_0 + \beta_1 \times \text{Both}_{imst} + \beta_2 \times \text{Captain Only}_{imst} \\ & + \beta_3 \times \text{All-Star Only}_{imst} + \gamma \times \mathbf{X}_{imst} + \alpha_{\text{Team}\times\text{Season}_{jt}} \\ & + \alpha_{\text{Opponent}\times\text{Season}_{kt}} + \alpha_{\text{Player Position}_i} + \varepsilon_{imst} \end{aligned} \tag{2.1}$$

Where:

- $\text{Performance}_{imst}$ represents the productivity of player i in match m during season s at time t.
- Both_{imst}, $\text{Captain Only}_{imst}$, and $\text{All-Star Only}_{imst}$ are dummy variables indicating the leadership status of player i: being both a captain and an All-Star, only a captain, or only an All-Star, respectively, in match m during season s at time t.
- $\mathbf{X}_{imst}$ includes control variables such as minutes played, player's age, and other game-specific factors that might influence a player's productivity.
- $\alpha_{\text{Team}\times\text{Season}_{jt}}$ and $\alpha_{\text{Opponent}\times\text{Season}_{kt}}$ are fixed effects that account for team-season and opponent-season interactions, capturing the influence of team dynamics and the competitive environment in each season.

- $\alpha_{\text{Player Position}_i}$ represents fixed effects for the player's position, controlling for the specific roles and responsibilities associated with each position on the court.

- ε_{imst} is the error term, capturing unobserved factors affecting player productivity. I assume that the error term is clustered at the player level to account for potential correlations within players across observations with the presence of fixed effects.

Basically, Plus-Minus measures the net point difference when a player is on the court, providing a metric of the contribution of players to the result while actively playing. However, this metric neglects the quality of teammates and opponents. To address this, Adjusted Plus-Minus was developed by employing statistical models to refine Plus-Minus. Adjusted Plus-Minus (APM) typically uses a regression approach that controls for dummy variables for each player on the court, both teammates and opponents, which allow it to estimate each player's impact while holding constant the effects of all other players. This approach helps to disentangle a player's individual contribution from the overall team performance. The details of APM can be found in the Appendix B.3. Building on this, Real Plus-Minus (RPM), developed by Stephen Ilardi who is an academic at Kansas University and an analyst at ESPN, incorporates additional player statistics and more sophisticated adjustments for team dynamics and opposition quality and offers a comprehensive metric that captures a player's overall impact with greater precision (Ghimire et al., 2020). While the exact methodology is proprietary, RPM is known to include several key components such as a box score prior to stabilising estimates, multi-year data to reduce noise, positional dummies for fairer comparisons, and regularisation techniques to prevent outlier results for players with limited playing time (Ghimire et al., 2020).

Table 2.1 *Descriptive Statistics of Player-Match Level Data*

	N	Mean	St. Dev.	Min	Max
Game-Related Statistics					
Season	674,837	2,011.418	5.617	2,002	2,021
Playoff (1 if a playoff game)	674,837	0.066	0.248	0	1
Home Game (1 if played at home)	674,837	0.000	0.000	0	0
Team Elo Rating	674,490	1,510.981	109.748	1,155.440	1,865.449
Team Payroll (in nominal USD)	674,554	82,235,975	27,911,210	33,458,932	178,980,766
Total Match Score	674,837	100.874	13.270	53	196
Player-Related Statistics					
All-Star (1 if played in All-Star game)	667,132	0.148	0.355	0	1
Captain (1 if player is team captain)	667,132	0.195	0.396	0	1
Both (Captain and All-Star)	667,132	0.101	0.301	0	1
Captain Only	667,132	0.095	0.293	0	1
All-Star Only	667,132	0.048	0.213	0	1
Draft Number	550,439	20.877	15.109	1	75
Salary (in nominal USD)	459,527	5,486,141	6,237,637	0	43,006,362
Age	643,362	27.138	4.292	18	45
Experience in NBA (in years)	643,293	5.935	4.084	0	23
Player-Game Statistics					
Minutes Played in Game	527,015	23.464	11.409	0	65
Real Plus-Minus (RPM)	344,855	0.000	10.740	-60	59
Field Goals Made	527,015	3.643	3.067	0	28
Field Goals Attempted	527,015	8.029	5.767	0	50
3-Point Field Goals Made	527,015	0.741	1.202	0	14
3-Point Field Goals Attempted	527,015	2.078	2.519	0	24
Free Throws Made	527,015	1.763	2.384	0	26
Free Throws Attempted	527,015	2.322	2.925	0	39
Offensive Rebounds	527,015	1.063	1.427	0	18
Defensive Rebounds	527,015	3.067	2.714	0	25
Rebounds	527,015	4.130	3.535	0	31
Assists	527,015	2.133	2.507	0	25
Steals	527,015	0.733	0.985	0	10
Blocks	527,015	0.473	0.884	0	12
Turnovers	527,015	1.332	1.408	0	12
Fouls	527,015	2.033	1.508	0	6
Points	527,015	9.790	8.157	0	81
Starter (1 if player was in starting 5)	674,837	0.375	0.484	0	1
Did not Play (1 if player did not play)	674,837	0.128	0.334	0	1

Note: Data Sources: ESPN & FiveThirtyEight.

2.3.2 Results and Discussion

Table 2.2 descriptively shows mean salary and Real Plus-Minus (RPM) by player roles. I use the salaries of players as a proxy for the talent they have. I use nominal values of salaries as I include team-season fixed effects which capture the impact of inflation. Table 2.3 reports the regression results. When the controls and a set of fixed effects are added to absorb or control for unobserved heterogeneity between player positions and time-invariant factors specific to teams, players who are both captains and All-Stars perform better than their teammates significantly and contribute to the team production positively. Conversely, players designated as captains without All-Star experience demonstrate a significant negative performance differential when compared to non-leaders. This may raise a problem with the captain assignment processes in teams. Although such players may still be contributing to the team success in different areas including team cohesion, team captains might be expected to perform well too. On the other side, the performance of All-Star players who have not been assigned as captains shows no significant deviation from that of their non-leader teammates.

Table 2.2 *Salary, RPM, Minutes, and RPM per Minute by Player Role*

Role	Mean Salary	Mean RPM	Mean Minutes	RPM per Minute
Regular Player	3,995,778	-0.3361376	20.65859	-0.0780085
Captain Only	9,425,869	-0.0220852	27.83307	-0.0295563
Allstar Only	8,688,743	1.4186897	26.65745	0.0185244
Both	16,943,390	2.5865464	33.42574	0.0752740

Note: Note: Data Source: ESPN. RPM: Real Plus-Minus. Salary values are in USD.

According to these findings, captains who experienced All-Star games may be a source of productivity spillover. Moreover, although only captains and only All-Stars do not perform better than regular players, they still may help others to perform better. Although being a better performer could affect others easily, this could not be necessary as they can altruistically support others using different tools than in-game performance. In the next section, the analysis will focus on examining the impact of a leader's absence on the productivity of other team members. This investigation aims to understand the extent to which leaders affect their peers if they can and to determine if performing better is necessary for having such an effect on others.

Table 2.3 *Leader Type and Performance*

Dependent Variable:	Real Plus-Minus		
Model:	(1)	(2)	(3)
Both (Captain & All-Star)	2.937***	0.5894**	0.4975***
	(0.2603)	(0.1888)	(0.1659)
Captain Only	0.2980	-0.2617*	-0.3795***
	(0.1811)	(0.1309)	(0.1146)
Allstar Only	1.705***	0.2289	0.1063
	(0.3205)	(0.2289)	(0.2007)
Playoff		-0.0773	-0.0490
		(0.0959)	(0.0925)
Minutes		0.1140***	0.1204***
		(0.0036)	(0.0037)
Salary (10 Million USD)		0.0848	0.0389
		(0.0832)	(0.0806)
Elo Difference		0.0167***	-0.0035***
		(0.0003)	(0.0004)
Age		0.2583*	0.2502***
		(0.1013)	(0.0918)
Age^2		-0.0041*	-0.0038**
		(0.0018)	(0.0016)
Fixed-effects			
Player Position	No	No	Yes
Team $\times$ Season	No	No	Yes
Opponent Team $\times$ Season	No	No	Yes
Observations	344,549	288,226	288,226
R^2	0.008	0.074	0.096

Note: Data Sources: ESPN & FiveThirtyEight. Player-level clustered robust standard errors in parentheses. The reduction in observations from Model (1) to Models (2) and (3) is due to missing data for salary and Elo rating variables. * p<0.1; ** p<0.05; *** p<0.01.

When I examine the sub-metrics of players for performance, I find that captains who played in an All-Star game get more defensive rebounds and steal more balls, suggesting that they spend more effort than others while defending and, meanwhile, surprisingly commit fewer fouls. On the offence, although they lose the ball more than others together with only All-Stars, they assist more than the rest. However, both captains and All-Stars suffer from more fouls as stopping them could be challenging. While all types of leaders attempt more to score, their success rate in three-pointers is slightly lower than their non-leader teammates. On the other side, these players often work longer by staying longer on the court. This could indicate a diminishing rate of marginal productivity, where performance may decline as the game progresses due to fatigue. Table B.3 and Table B.4 in the Appendix report regression results of each box score metric.

2.4 Absence of Leaders

In this main part of the study, I use in-game injuries of leader players as a source for random and unexpected exogenous variation to control their effect on other players (J. S. Chen & Garg, 2018; Wegelin et al., 2022). Throughout the games, I keep track of the players who are actively playing on the court and their leadership status (whether they are team captains or All-Stars) using the in-game running lists of players provided for each team. Therefore, I ensure that the injured leader player is substituted with a non-leader to disentangle the effect of leaders on others. Previous studies tried to infer players on the court using event data in the absence of in-game running lists of players and because of that reason needed to drop several observations (Arcidiacono et al., 2017). Furthermore, I can control for the playing time of players which can be seen endogenous because of the variation in productivity of players in games.

2.4.1 Data and Methodology

In this part, I use play-by-play data from NBA seasons from 2002 to 2021. Play-by-play data have been obtained from ESPN while injuries data have been obtained from Pro Sports Transactions, running lists of active playing players in games from Basketball-Reference and ELO ratings from FiveThirtyEight. As there are not any data for in-game injuries available, I use a similar technique to Wegelin et al. (2022) to find such games. First, I find games where leaders were absent because of an injury. Then, I checked the last game they played before they were reported as injured. If a leader player leaves the game before the end of the third quarter and never plays anymore, I record it as an in-game injury. Additionally, I gradually relaxed this assumption to include players leaving games until the last two minutes, and the findings remain robust. The descriptive statistics of play-by-play data are provided in Table 2.4.

Table 2.4 *Play-by-Play Descriptive Statistics*

Statistic	N	Mean	St. Dev.	Min	Max
Game-Related Statistics					
Home (1 if game played at home)	10,758,743	0.502	0.500	0	1
Playoff (1 if a playoff game)	10,758,743	0.066	0.249	0	1
All-Star (1 if played in all-star game)	10,758,743	0.232	0.422	0	1
Captain (1 if team captain)	10,758,743	0.291	0.454	0	1
Both (1 if both captain and all-star)	10,758,743	0.178	0.382	0	1
Home Score	10,758,743	52.884	30.999	0	168
Away Score	10,758,743	51.153	30.230	0	168
Elo Rating	10,754,902	1511.127	110.130	1155.440	1865.449
Team Payroll (in USD)	10,758,743	84,282,942	28,490,468	33,458,932	178,980,766
Scoreline (score difference at point)	10,758,743	1.731	10.432	-78	78
Player-Related Statistics					
Player Salary (in USD)	7,081,047	7,886,410	7,610,556	0	43,006,362
Age	7,600,027	26.654	4.103	18	44
Experience (in years)	7,593,696	6.028	3.995	0	23
Draft Number	6,901,023	18.006	14.612	1	75
Play-Related Statistics					
Scoring Play (1 if scored)	10,758,743	0.251	0.433	0	1
Score Value	10,758,743	0.425	0.850	0	3
Shooting Play (1 if a shot attempted)	10,758,743	0.480	0.500	0	1
Half	10,758,743	1.509	0.500	1	2
Period/Quarter	10,758,743	2.549	1.139	1	8
Distance (in feet)	10,758,743	16.753	10.263	0	93
Cumulative Fouls	10,044,537	1.119	1.180	0	6
Play-Related Dummies					
Free Throw	10,758,743	0.015	0.120	0	1
Two Point	10,758,743	0.233	0.423	0	1
Three Point	10,758,743	0.003	0.053	0	1
Foul	10,758,743	0.093	0.291	0	1
Ejection	10,758,743	0.00001	0.003	0	1
Turnover	10,758,743	0.060	0.237	0	1
Rebound	10,758,743	0.229	0.420	0	1
Dunk	10,758,743	0.016	0.126	0	1
Layup	10,758,743	0.080	0.271	0	1
Absence-Related Statistics					
Allstar Injury Dropout	10,758,743	0.004	0.066	0	1
Captain Injury Dropout	10,758,743	0.006	0.079	0	1
Allstar 6^{th} Foul	10,758,743	0.0002	0.015	0	1
Captain 6^{th} Foul	10,758,743	0.0003	0.017	0	1

Note: Data sources: ESPN & FiveThirtyEight

Unfortunately, Real Plus-Minus (RPM) data are not provided for play-by-play data. To be able to capture the team productivity and individual contribution before and after the treatment, I trained long short-term memory (LSTM) networks using the historical performance of players embedded in the play-by-play and end-of-game RPM of players. LSTM networks are a kind of recurrent neural network (RNN) system designed to learn from sequences of data by capturing important patterns over long intervals and are highly used to predict data based on time series (Hochreiter & Schmidhuber, 1997). I chose LSTMs over other types of RNN because of their ability to capture long-term dependencies in sequential data which is crucial for modeling the changing dynamics of player performance throughout a game. The architecture of LSTMs allows model to learn which information is relevant to keep or forget over short and long timescales. This is important in sports, especially in basketball, where an early play may effect strategy of later plays in the game or where a player's performance may follow patterns across multiple games. Moreover, play-by-play data are inherently noisy as there are other outcomes than scores such as fouls, turnovers and timeouts. LSTMs provide robust performance in handling such noisy sequential data (Greff et al., 2017). Their ability to filter out irrelevant information while retaining important patterns helps in generating more accurate predictions of RPM throughout the game.

The model predicts the RPM of players at every point and shows any changes from the previous one to the next in the play-by-play data in sequential order. LSTM architecture processes the sequence of game events such as player actions in the current play, action times and scoreline in this play. At each time step (each play), the model predicts an RPM value for each player on the court by using Player IDs. This approach allows me to capture the temporal dynamics of player performance throughout the whole game. The LSTM model was initially trained on the first 16 seasons in the dataset by using an approximately 80-20 split for training and validation. As this is a sequential data, the training data were not chosen randomly. Then, I validated the end-of-game RPM values of the trained model with that of ESPN by using the data from the remaining 4 seasons.

Detailed LSTM specifications and Figure B.1 which shows predicted vs ESPN RPM, are provided in the Appendix B.4.

I then use the predicted Real Plus-Minus (RPM) variable to estimate the impact of a leader's absence (the treatment) on the productivity of the player who takes the final shot in a possession. Although one might argue that this variable should only be used for the shooting player in each play as an observation, this approach is appropriate for measuring productivity in basketball. This is because the outcome of a possession largely depends on the actions and success of the player who last handles the ball. By focusing on the final shooter, the model captures the effect of a leader's absence, as the leader could have provided an assist or created better scoring opportunities for their teammates had they been on the court.

It is important to acknowledge that APM may not directly control for leadership qualities. As Arcidiacono et al. (2017) argues, APM can capture both direct productivity and spillover effects of players. Consequently, using APM in peer effects analysis could be challenging as findings could be biased. However, RPM may perform better in this regard. By incorporating players' historical performances, RPM could help to focus more on a player's individual contributions, potentially disentangling some of the spillover effects, though not entirely.

Furthermore, the use of predicted RPM in this study offers additional mitigation for potential estimation problems associated with APM and RPM metrics. Unlike RPM, which automatically adjusts for lineup changes, the predicted RPM is based on end-of-game data. LSTM networks provide mitigation by capturing long-term dependencies and complex patterns in sequential and historical game data. This approach reduces the potential for immediate adjustments in RPM due to player substitutions, including those caused by leader injuries. However, it is important to note that the predicted RPM may still indirectly reflect some teammate effects, as it is trained on RPM which captures such effects. This issue creates a potential bias towards underestimating the effect of leaders on the performance of other players.

Despite this potential underestimation, the study still finds significant effects of leaders. This underscores the robustness of the findings. The use of LSTM networks to predict RPM at each point in the game provides a more robust measure of performance that is limitedly sensitive to immediate adjustments following leader absences.

In the main analysis, I use staggered difference-in-differences event study (Callaway & Sant'Anna, 2021), which enables me to detect a precise treatment effect of injuries which take place at different times during the games. This approach offers several advantages over OLS. While OLS estimates would likely be biased due to endogeneity issues stemming from unobserved factors such as game dynamics and player fatigue, staggered DiD addresses these concerns more effectively. By exploiting the precise timing of in-game injuries as exogenous shocks, staggered DiD can isolate the effect of leader absence from other confounding factors. Moreover, reverse causality might exist if poor team performance increases the risk of injuries. Some players may avoid taking responsibility and feign injury if a loss is unlikely in a game. These issues make causal inference challenging with OLS. The staggered DiD method addresses these concerns by exploiting the exogenous nature of in-game injuries and controlling for time-invariant confounders. Unlike OLS, which assumes a constant treatment effect, staggered DiD allows for heterogeneous treatment effects across different time periods which captures the effects of leaders in a game.

Previous studies, such as Wegelin et al. (2022), which examined the effect of high-performing players on their teammates in the NBA using injuries as treatment, employed a canonical difference-in-differences approach with two groups and two periods. However, this method may lead to misleading results if the treatment effect is heterogeneous across different groups or over time (de Chaisemartin & D'Haultfœuille, 2023). Essentially, the effect of an injury in the last minutes might be different than an earlier one. Besides the event study, I estimated the treatment effect using canonical difference-in-differences as a robustness check while estimating treatment effects separately for each leader type and controlling for players on the court.

Besides in-game injuries, I also use fouled-outs of leader players in a separate analysis. Although committing 6^{th} foul may not be random and unexpected, it is still worth analysing how non-leaders react in games. In addition to performance metrics, I use coordinates of in-game events on the court and use them to understand if teams struggle more to get closer to the basket to score less risky shots in the absence of creative leaders. Such a risk affects the productivity of players shooting successfully is more challenging when the distance is bigger. This idea was suggested by Wegelin et al. (2022) but was not tested because of the lack of data. By using play-by-play data coordinates, I calculate the distance of shots to the basket and test if teams are forced to shoot from away instead of promising positions when leaders are absent. Table B.5 in the Appendix shows the distribution of pre/post-control and treatment observations.

In the next part, I will analyse the effect of leader in-game dropouts on the performance of non-leaders in the team and overall team productivity, separately.

2.4.2 Effect of Absence of Leader on Performance of Other Players

In this part, I analyse the effect of leader absence on the performance and productivity of non-leaders using RPM and box score metrics. I begin with the tests of parallel trends using placebo treatment times $< n$ where n is the actual treatment time, expecting an insignificant coefficient of interest and analysis shows that the parallel trends assumption holds. Then, the average treatment effect on the treated is estimated using Callaway & Sant'Anna (2021) difference-in-differences and the model below for the absence of each leader type. A set of players' fixed effects is added to absorb the differences in the leadership status of players besides their other characteristics on the court besides game, team and opponent fixed effects. Robust standard errors are clustered at game and player levels.

$$
\begin{aligned}
\text{Performance}_{itp} = \beta_0 + \beta_1 \times \text{LeaderAbsence}_{tp} + \gamma \times \mathbf{X}_{itp} + \alpha_{\text{Player}i} + \alpha_{\text{Game}t} \\
+ \sum_{j=1}^{3} \alpha_{\text{Teammate}_j} + \sum_{k=1}^{5} \alpha_{\text{Opponent}_k} + \varepsilon_{itp}
\end{aligned}
\tag{2.2}
$$

Where:

- Performance_{itp} denotes the performance metric of player i at play p in game t, reflecting the real-time productivity of non-leader players.
- $\text{LeaderAbsence}_{tp}$ is a dummy variable equal to 1 if a leader is suddenly absent (due to injury or fouling out) and replaced by a non-leader in play p in game t for player i and 0 otherwise.
- $\mathbf{X}_{itp}$ includes play-specific control variables that might influence performance, such as the current score difference, time remaining in the game, and the quarter.
- $\alpha_{\text{Player}i}$ represents individual player fixed effects, accounting for unobserved characteristics of player i that could affect player productivity.
- $\alpha_{\text{Game}t}$ captures game-specific fixed effects, reflecting characteristics of game t that could affect player productivity.
- $\sum_{j=1}^{3} \alpha_{\text{Teammate}_j}$ and $\sum_{k=1}^{5} \alpha_{\text{Opponent}_k}$ are the sums of fixed effects for the three teammates (except player i and leader who will be injured or fouled out) and five opponents on the court, respectively, during play p, controlling for the influence of other players type (e.g. any changes in the number of leaders of opponent) in the game.
- ε_{itp} is the error term for player i at play p in the game t, capturing unobserved factors that might affect performance during that specific play.

Moreover, along with Callaway & Sant'Anna (2021) difference-in-differences, I estimate canonical difference-in-differences and provide its results in the Results and Discussion part.

2.4.3 Effect of Absence of Leader on Team Success

Finally, in this part, I estimate the effect of leader absence on team productivity using game outcome (win/loss) and score difference at the end of games as dependent variables and logistic and OLS regression, respectively.

I estimate the effect of the absence of leaders on team productivity using the equation given below.

$$\begin{aligned}\log\left(\frac{P(\text{Win}_{jkt}=1)}{1-P(\text{Win}_{jkt}=1)}\right) &= \beta_0 + \beta_1 \times \text{AllStar}_{jt} + \beta_2 \times \text{Both}_{jt} + \beta_3 \times \text{Captain}_{jt} \\ &\quad + \gamma \times \mathbf{X}_{jkt} + \alpha_{\text{Team}\times\text{Season}_{jt}} + \alpha_{\text{Opponent}\times\text{Season}_{kt}} + \varepsilon_{jkt}\end{aligned} \tag{2.3}$$

Where:

- $P(\text{Win}_{jkt}=1)$ is the binary outcome of the game between team j and opponent k at time t, where 1 represents a win and 0 a loss.
- AllStar_{jt}, Both_{jt}, and Captain_{jt} are dummy variables indicating the injury of team j's leaders (All-Star only, both All-Star and Captain or Captain only) at time t.
- $\mathbf{X}_{jkt}$ represents control variables such as absence leaders of the opponent team, playoff dummy, Elo rating differences, and other factors relevant to the game outcome.
- $\alpha_{\text{Team}\times\text{Season}_{jt}}$ and $\alpha_{\text{Opponent}\times\text{Season}_{kt}}$ are fixed effects for interactions between team j and season at time t and opponent k and season at time t, respectively.
- ε_{jkt} is the error term for the game between team j and opponent k at time t.

The results of these estimations are provided in Table 2.7 along with their discussions in the Results and Discussion part below.

2.4.4 Results and Discussion

Table 2.5 reports the results of canonical difference-in-differences analysis, which estimates the effect of injuries and fouled-outs of leaders on the performance of other players. This 2×2 difference-in-difference approach shows that the absence of both(captain and All-Star) due to injuries decreases the performance of other players. The absence of captains or All-Stars and the absence of both due to the 6^{th} fouls do not significantly affect the performance of other players.

This could be because the timing of the 6^{th} foul could be predictable as some players have this problem very often. Therefore, teams can develop long-run strategies to mitigate such impacts. While injury absences are unpredictable, teams are not ready for it and the effect is more obvious. Exposure to treatment (absence of leader) could be longer in injuries as committing six fouls takes time while injuries could occur at any time, instantaneously. Similarly, in businesses, organisations might be more prepared for predictable absences of workers and can be prepared for it while unexpected ones could cause some crisis.

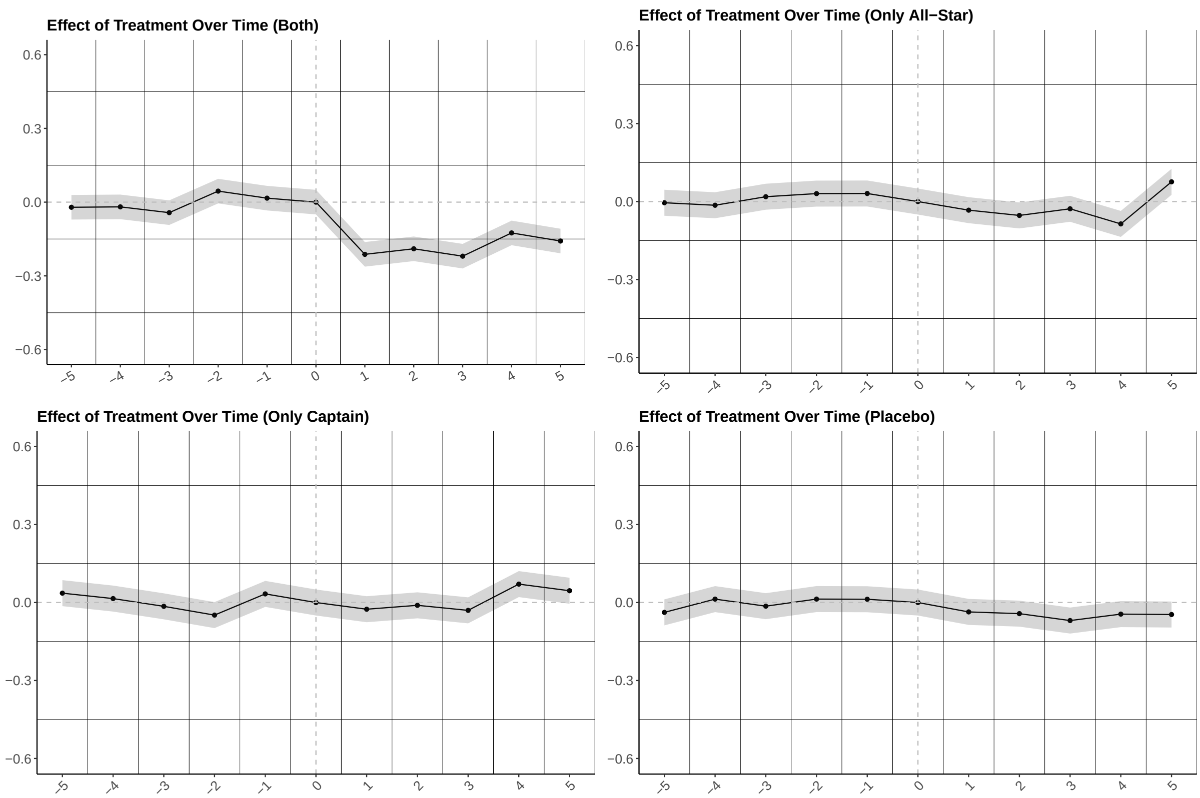
Table 2.5 *Effect of Leader Absence on Performance: Canonical Difference-in-Differences Analysis*

Dependent Variable:	*Real Plus-Minus*					
Treatment Reason:	Injury	6^{th} Foul	Injury	6^{th} Foul	Injury	6^{th} Foul
Treatment $\times Post : Both$	-0.77 (0.31)**	-0.16 (0.10)				
Treatment $\times Post : OnlyAll - Star$			-0.52 (0.31)	-0.11 (0.07)		
Treatment $\times Post : OnlyCaptain$					-0.30 (0.21)	-0.07 (0.07)
Play-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects						
Game	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Teammates	Yes	Yes	Yes	Yes	Yes	Yes
Opponent Players	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,752,100	4,752,100	4,752,100	4,752,100	4,752,100	4,752,100
R^2	0.34	0.34	0.34	0.34	0.34	0.34

Data source: ESPN & FiveThirtyEight. Play-level controls include remaining time for the end of the quarter, half-time and match, and scoreline. Two-way (Game & Player-level) clustered robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

Figure 2.1 illustrates the RPM of non-leaders before and after the injury of players by the player group along with the placebo leaders. Similar to the findings of canonical difference-in-differences, the result of the event study shows that the RPM of non-leaders is negatively affected by the injury of players who are both team captains and All-Stars. However, injuries of only captains and only All-Star players have no significant effect on the performance of non-leaders. To verify if these effects are special to such players or if it is a general mechanical effect, I conducted a placebo test. I randomly selected a group of non-leader players who got injured at some point in the game and treated them as if they were leaders. Then, I replicated the analyses above using the injuries of these placebo leaders as an exogenous shock. However, the results of the placebo tests reinforced the findings, indicating that the observed effects are not merely mechanical but are specific to the absence of actual leaders.

Fig. 2.1 *Real Plus-Minus (RPM) of Regular Players Before and After Injury of Player Type*



Note: Data source: ESPN. Callaway & Sant'Anna (2021) event study estimations before and after treatments. X-axes: Time to Treatment (in Minutes). The shaded areas represent 95% *confidence intervals.*

As shown, team captains who never played in an All-Star game and players who played in an All-Star game but do not have the team captaincy, do not have any effect on the performance of other players. The reason for that could be that they are not better than regular players in teams. This shows that to be able to affect others, leaders need to be a source of a spillover. Some players could be chosen for the All-Star game or assigned as captains because of their popularity. However, the findings show that talent is vital for leading others. A person could be a superstar either through talent (Rosen, 1981) and/or through popularity (Adler, 1985). However, being an effective leader requires superstardom as defined by Rosen (1981).

This finding aligns with that of Wegelin et al. (2022), which reports a decrease in the field goal percentage of non-high-performing players following the injury of a high-performing teammate. It is conceivable that the high-performing players referred to in their study could correspond to the All-Starred team captains discussed in this analysis as both perform better than other players and have a similar impact on the performance of their teammates and team. However, contrary to Wegelin et al. (2022), my analysis reveals that teams attempt more three-pointers rather than two-pointers in the absence of players who hold both captain and All-Star status. When such leaders are present on the court, 74.4% of shots (combining two and three-point attempts) are two-pointers; this proportion drops to 60.8% in their absence. This shift may derive from the lack of creativity typically contributed by the All-Starred team captain.

Further, by analysing the location and distance of shots, I find that three-point field goals are attempted from greater distances when only All-Starred team captains are absent, whereas the distance of two-point field goals remains unaffected. This could be attributed to the more constrained area for two-pointers and teams opting for two-point shots only when good opportunities arise as the proportion of two-pointer attempts goes down. Also, the effect of the absence of only All-Star players is significant under 10% level showing weak importance of their talent and creativity on the distance of three-pointers. Table 2.6 below shows the effect of the absence of leaders on the distances of three-pointers using difference-in-differences estimation. Table B.6 shows a similar table for two-pointers in the Appendix.

Table 2.6 *Distance of Three-Point Field Goal Attempts and Absence of Leaders*

Dependent Variable:	Distance of Three-Point Field Goal Attempts		
Model:	(1)	(2)	(3)
Both: Treatment × Post	0.5580***		
	(0.2098)		
Captain: Treatment × Post		0.3474	
		(0.2294)	
All-Star: Treatment × Post			0.4400*
			(0.2592)
Home	-0.0654***	-0.0657***	-0.0656***
	(0.0094)	(0.0093)	(0.0093)
Score Difference	0.0021***	0.0021***	0.0021***
	(0.0006)	(0.0006)	(0.0006)
Period	-0.0902***	-0.0898***	-0.0888***
	(0.0067)	(0.0067)	(0.0067)
Fixed-effects			
Game	Yes	Yes	Yes
Player	Yes	Yes	Yes
Team	Yes	Yes	Yes
Opponent Team	Yes	Yes	Yes
Observations	683,102	683,102	683,102
R^2	0.110	0.110	0.110

Note: Data source: ESPN. Score Difference is *Team – Opponent*. Player-level clustered robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

Furthermore, Table 2.7 reports the logistic regression results of the game result and OLS regression results of score difference at the end of the game with the absence of leader types for home teams only to avoid correlated error terms of duplicate observations. When a player who is a team captain and experienced an All-Star game, is absent, the chance of the team winning the game decreases around 24% $(1 - e^{-0.27})$ for home teams and approximately 20% $(1 - e^{-0.22})$ for away teams when controls and fixed effects are included. Moreover, in the absence of such players score difference significantly decreases even if teams win the game, meaning teams struggle more. Therefore, team productivity is significantly lower when such key players are missing. For a robustness check, I replicated the analyses using observations of teams that had only one player from any player type and their absence due to injury and found similar results. Table B.7 in the Appendix shows the regression results for away teams.

When the findings of the study are considered, Tables 2.3 and 2.5 and Figure 2.1 reveal key insights about leadership in basketball teams. Table 2.3 shows All-Star players perform better than non-leaders when not accounting for age and salary. However, Table 2.5 and Figure 2.1 indicate that the absence of only All-Stars cannot affect their teammates' performance significantly while when All-Stars are team captains, their absence negatively affects other players' performance. The absence of players who are only captains does not have this effect. This suggests that being a skilled player alone is not enough to influence team performance.

These findings show that there is a behavioural effect beyond individual talent. Both All-Star captains seem to have a unique combination of expertise and formal leadership that allows them to affect their teammates positively. This study shows evidence that these players are better leaders who are able to enhance their teammates' performance beyond their own contributions on the court.

Table 2.7 *Score Difference, Game Result and Injury of Key Players (Home Teams)*

Dependent Variables:	Score (1)	Score (2)	Score (3)	Result (1)	Result (2)	Result (3)
Injury of Both	-1.57^{***}	-1.23^{***}	-1.89^{***}	-0.19^{***}	-0.16^{***}	-0.27^{***}
	(0.20)	(0.18)	(0.25)	(0.03)	(0.03)	(0.05)
Injury of Only Captain	-1.59^{***}	-0.30	-0.17	-0.20^{***}	-0.02	0.01
	(0.17)	(0.16)	(0.25)	(0.03)	(0.03)	(0.05)
Injury of Only All-Star	0.15	-0.27	-0.54	0.04	-0.01	-0.08
	(0.19)	(0.17)	(0.28)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Both	1.20^{***}	0.87^{***}	1.65^{***}	0.13^{***}	0.10^{**}	0.22^{***}
	(0.20)	(0.19)	(0.25)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Only Captain	1.28^{***}	0.09	0.57^{**}	0.18^{***}	0.01	0.08
	(0.17)	(0.15)	(0.24)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Only All-Star	-0.64^{***}	-0.17	0.12	-0.07^{*}	-0.00	0.05
	(0.19)	(0.17)	(0.28)	(0.03)	(0.03)	(0.05)
Play-off		1.22^{***}	0.74^{*}		0.11^{*}	0.03
		(0.32)	(0.35)		(0.05)	(0.06)
Elo Difference		0.04^{***}	-0.01^{***}		0.01^{***}	-0.00^{***}
		(0.00)	(0.00)		(0.00)	(0.00)
Fixed-effects						
Team × Season	No	No	Yes	No	No	Yes
Opponent Team × Season	No	No	Yes	No	No	Yes
Observations	25562	25549	25549	25562	25549	25549
R^2	0.01	0.16	0.25			
Deviance				34368.13	31001.53	28727.38
Log Likelihood				-17184.06	-15500.76	-14363.69
Pseudo R^2				0.01	0.10	0.10

Note: Data source: ESPN & FiveThirtyEight. The first three columns show OLS estimations where the dependent variable is the score difference. The remaining three columns show logistic regression where the dependent variable is the match result. Game-level clustered robust standard errors in parentheses. The decrease in observations from Models (1) and first models to the other models is due to missing data for the Elo Difference variable. * p<0.1; ** p<0.05; *** p<0.01.

The findings of this study contribute to the broader literature on peer effects by providing evidence of positive peer effects in a high-stakes team environment. While some studies have found negative or negligible peer effects, such as Waldinger (2012) which finds no significant peer effects among scientists, and Brady et al. (2017) which demonstrates negative peer effects on student achievements, this study aligns with research showing positive peer effects. The significant positive impact of players who are both team captains and All-Stars on their teammates' performance is consistent with findings by Mas & Moretti (2009) and Bandiera et al. (2010) which demonstrate positive productivity spillovers from high-performing workers to their peers in different workplace settings.

Moreover, this study's results are in line with peer effects literature in sports that find positive effects. For instance, the findings are in line with Arcidiacono et al. (2017) and Wegelin et al. (2022) which also find positive spillover effects from high-performing basketball players to their teammates. Similarly, Gould & Winter (2009) and Jane (2015) show positive peer effects in baseball and swimming, respectively. This study extends these findings by demonstrating that the positive peer effects come from individuals who combine formal leadership roles (team captains) with recognised exceptional skill (All-Stars) by suggesting that the mechanism of peer effects in team sports may be more complex than simply skill or authority alone.

2.5 Concluding Remarks

In this study, I examined if the unexpected loss of a talented leader worker affects the productivity of others in teams using data from the NBA. The event study shows that the absence of players who have both captaincy and star-title altogether negatively affects the productivity of individuals and teams while the absence of players assuming only captaincy or star title neither affects the performance of other players nor that of teams. The findings imply there could be some inefficiencies in assigning team captains.

Although they may have intangible effects on teams including team cohesion and social leadership, motivating others through performance could be expected too. Teams can benefit from the role of the team captain to identify them as an idol or role model in teams that promote an intrinsic motivation to perform for team identity. Although being very talented is beneficial for team production, only star players are not able to be a source of spillover and affect others. I conclude that for a peer effects to be significant, certain conditions must be met, such as holding a formal leadership position and possessing superior talent. Without these attributes, the presence or absence of players and managers may not substantially affect the performance of others in teams.

Additionally, by using the distance of shoots, I showed that teams need to take more risks when leaders are absent. That could be the case for employees who need to make decisions affecting their firms and others such as sales and purchasing departments. Managers would have negotiation experience in such transactions and, therefore take fewer risks. However, their absence could be more costly in such operations. This finding shows the importance of both formal and informal leadership roles besides individual skills in driving team success and highlights the complexity of peer dynamics within teams.

The findings on leadership in basketball teams offer insights applicable to broader workplace dynamics. The finding that only players who are both team captains and All-Stars significantly impact their teammates' performance highlights the importance of combining formal authority with recognised expertise. This result may parallel potential workplace scenarios where the most effective leaders might be ones who have official management positions and are highly skilled in their field at the same time. By examining these effects in the high-pressure, quantifiable environment of professional basketball, this study provides a perspective on leadership effectiveness. These insights could inform how organisations in various industries select and develop their leaders by suggesting that the most influential figures may be those who blend formal authority with expertise in their role.

Limitations and Future Work

While this study sheds light on the impact of leader absence on team productivity within the context of the NBA, the findings are subject to limitations which could be addressed in future research. The main limitation is the study focuses on a single sport, which may limit the generalisability of the results to other sports or team-based settings. Additionally, the use of NBA play-by-play data, may not be capturing all aspects of leadership and team dynamics that affect the performance of individuals. Future studies could expand this research by replicating the analysis using data from other sports, such as football, hockey, or baseball, to examine whether the observed effects of leader absence hold in different team environments. Moreover, extending the investigation outside of sports, such as teams from other industries, could provide valuable insights into the role of leaders in team productivity. Such cross-disciplinary research could enhance our understanding of how the absence of key individuals affects team outcomes and inform strategies for mitigating these negative effects.

Chapter 3

The Effect of Patience on Tenure of Managers: Evidence From Football

> "On a long enough timeline, the survival rate for everyone drops to zero."
> — Chuck Palahniuk, *Fight Club*, 1996

3.1 Introduction

If Sir Alex Ferguson had embarked on his managerial career in Turkey, would he have been able to maintain his position for nearly three decades, as he did at Manchester United? This question may seem speculative at first glance; however, it serves as an entry point into the broader discussion of how the patience level of societies affects the tenure of employment, particularly in high-pressure roles like football managers (head coaches) or company CEOs.

Sports, especially football, are more than just games; they can reflect the societal and cultural norms of the countries where it is played. Football managers can be seen as very similar to company CEOs in other businesses. CEOs are formally responsible for the performance of their organisations as football managers are responsible for the performance of their teams. Also, CEOs and football managers reach their positions at similar ages when they gain experience in other positions and earn comparable amounts which are likely to exceed millions of pounds besides performance-related bonuses.

In football, each team is led by a manager who is also known as a head coach in some countries. A manager's role includes many responsibilities. These include selecting the team lineup, deciding on tactics and game plans, making substitutions during matches, and overseeing player training and development. Managers are also often involved in choosing which players to buy or sell. The performance of a team is closely linked to the effectiveness of its manager. Because of this, managers face constant pressure from fans, media, club boards and owners. Their job security can be uncertain and they are often dismissed if the team performs worse than expected. A manager's success is typically measured by the team's position in the league table, how many points they collect, and performance in cup competitions.

Despite these similarities, football provides data on the personal characteristics and performance of managers while such data are hard to obtain and analyse in other businesses where CEOs work. On the other side, objectively evaluating the performance of CEOs is challenging as there are several factors which may affect the performance of firms. However, measuring and evaluating the performance of football managers objectively is easier according to the expectations and the results.

In this study, I use data from football to understand how a socioeconomic factor, namely the patience of a country, affects the employment tenure of top-level managers. Alex Ferguson is an extreme example. His long tenure at Manchester United was sustained not just by his exceptional managerial skills but also by a society that values long-term planning, consistency, and a tolerance for unsatisfying results especially during the first couple of years in the position.[1] However, then he became one of the most successful football managers in history thanks to the patience of fans and owners of the club. Had

[1] Manchester United became 11$^{\text{th}}$ and 12$^{\text{th}}$ in the league in the first and third seasons of Ferguson and could not win the league until the seventh season in the team. In his third and fourth seasons, United could only win a third of the matches played.

the club decided to dismiss Ferguson after a few disappointing seasons, he might not have had the opportunity to achieve such remarkable success, even with his exceptional talent.

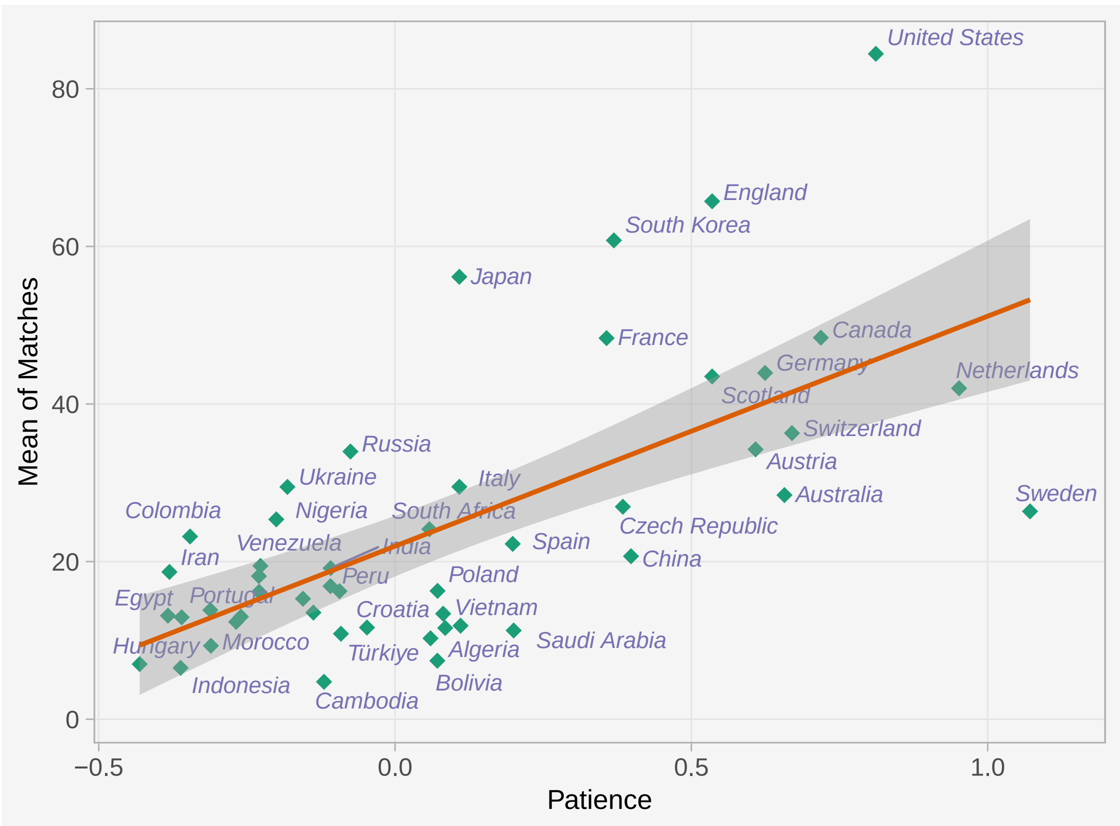
Similarly, highly skilled CEOs may be fired during recessions, even though the economic downturn, rather than their leadership capabilities, is the primary cause of the company's poor performance. Without patience and understanding from stakeholders, even legendary managers like Ferguson would not have the chance to prove their worth and leave a lasting impact on their teams and organisations. This anecdotal example is in line with findings of Blank & Hadley (2021) which shows that CEOs who experienced a recession in their tenure, perform better which leads to higher firm value. Similarly, Manchester United under Alex Ferguson became the most valuable sports team between 2010 and 2012 according to Forbes (2012, 2010, 2011) The World's 50 Most Valuable Sports Teams.

It might be argued that in societies where patience is a valued commodity, individuals in various roles, from executives to sports coaches, may enjoy longer tenures. This patience may arise from cultural, economic, and social factors. In contrast, societies characterised by a higher turnover in such roles often show an underlying impatience or a demand for rapid success which are influenced by various factors including economic pressures, cultural norms, or public expectations.

The dismissal of managers in football clubs can often be influenced by various theories. One prominent theory is the scapegoating theory, which suggests that managers are sometimes fired as a way for the club to deflect blame and appease frustrated stakeholders, even if the manager's performance is not the primary reason for the team's poor results (Grusky, 1963). This theory might be particularly relevant in impatient societies, where the pressure for immediate success is high, and tolerance for underperformance is low. In countries where the patience level is low, football clubs could be more likely to engage in scapegoating behaviour, as they face high pressure to deliver good results quickly. This pressure can come from various stakeholders, including fans, media and investors, who have high expectations and little tolerance for failure. In such environments, clubs may see dismissing managers as a symbolic act to demonstrate that they are taking action, even if the manager is not solely responsible for the team's performance.

It is worth noting that such pressure arising from the impatience of societies causes job insecurity which also leads to low performance and commitment in workplaces (Anand et al., 2023). Similarly, Slade & Tolhurst (2019) theoretically and empirically shows that NFL coaches take more risks when their job security is low. Their findings could be held in other sports and businesses too. Pressure stemming from the impatience of societies may force coaches to take unnecessary risks to succeed quickly which may lead to failure in competitions. Therefore, managers in impatient countries can find themselves in a fight club to survive; however, their timeline cannot be as long as their colleagues in patient countries. Figure 3.1 shows the correlation between country-level patience and the average number of matches coached by managers during their tenures.

Fig. 3.1 *Patience and Mean Number of Matches Coached by Managers by Countries*



Note: The measure of patience is from Falk et al. (2018). The line illustrates the correlation between the patience levels of countries and the average number of matches coached by managers within those countries. The grey shaded area represents the 95% confidence interval.

This study examines the relationship between the patience levels of countries and the tenure time of managers by using data from football. By examining the tenure of football managers across various countries, this research not only offers insights into employment practices in sports but also tries to provide a broader understanding of how cultural and socioeconomic factors shape employment practices in various sectors.

The hypothesised causal connection between the patience levels of countries and football manager tenure operates through several mechanisms. While football fans may not be a representative sample of the broader society, their patience levels could correlate with societal norms. A country's patience is likely transmitted through social learning and institutions by influencing the expectations and behaviours of individuals. This includes fans, whose patience levels are expected to broadly reflect societal norms, though individual variations can exist. In more patient societies, fans may be more willing to tolerate periods of underperformance by believing in long-term strategies.

On the other side, the media, which is influential in shaping public opinion, may also reflect these societal patience levels in their narratives about manager performance. Furthermore, club boards, operating within this cultural context, may be more inclined to give managers time to implement their strategies in more patient societies. While I use country-level patience as a proxy for these mechanisms, I also acknowledge that club-level variations in patience can exist. The relationship between a country's patience and that of a club's fans, boards and owners is not necessarily the same but could be highly correlated.

The analysis shows that the patience level of a country and the tenure durations of football managers are positively correlated. In order to examine if this correlation implies a causal link, I use two different instrumental variables for patience and find identical results. Moreover, a survival analysis, namely the accelerated failure time model, confirms the findings.

3.2 Literature Review

The literature review is structured into two main parts. Initially, the focus will be on studies examining patience and its impact within the field of economics. Then, the literature on football manager/coach dismissals will be reviewed in the second section.

3.2.1 Patience

Although patience could play a crucial role in various aspects of society, including economic growth, educational attainment, and individual well-being, measuring the patience levels of societies is a challenging task, as it involves capturing a complex and multifaceted concept. However, the Global Preferences Survey (GPS) of Falk et al. (2018) provides a comprehensive dataset from 76 countries to study patience. This survey defines patience as the willingness to wait for a bigger reward rather than getting a smaller reward today. The researchers used experimentally validated survey questions to capture preferences related to patience. Their findings revealed substantial variations in patience across countries and regions. For instance, the most patient countries were in Western Europe, such as Sweden, Denmark, and Germany, while several Asian and African countries exhibited lower levels of patience. The study found that patience is positively associated with average years of schooling, gross national savings, and GDP.

Before the study of Falk et al. (2018), researchers used long-term orientation (LTO) of Hofstede (2001) as a proxy for patience although survey items of LTO are distant from time and risk preferences (Falk et al., 2018). Also, it is claimed that the measure of LTO is not representative of societies as it was surveyed with employees of a company, IBM, only (Falk et al., 2018).

In addition to the GPS, another dataset that includes patience measures is an international survey conducted by Wang et al. (2016). This study collected data from 53 countries and found that all countries exhibit hyperbolic discounting patterns, with the immediate future being discounted more heavily than the distant future. The study also reveals that cultural factors, as captured by Hofstede (2001)'s cultural dimensions, significantly contribute to the variation in time discounting across countries. Specifically, higher degrees of Individualism and Long-Term Orientation (LTO) predicted a stronger tendency to wait (patience) for larger payoffs.

Furthermore, Wang et al. (2016) finds that patience is correlated with several important factors. Countries with higher levels of patience tend to have higher rates of innovation, as measured by the number of patents per capita. This suggests that patient societies may be more likely to invest in long-term projects and research, leading to technological advancements. Patience is also positively correlated with environmental protection, indicating that patient countries may be more willing to make short-term sacrifices for long-term environmental benefits. Additionally, patient countries tend to have higher credit ratings, which could be attributed to their ability to manage debt and make long-term financial decisions. Finally, the study finds a negative correlation between patience and body mass index, suggesting that individuals in patient societies may be more likely to make healthier lifestyle choices, such as maintaining a balanced diet and engaging in regular exercise.

The relationship between patience, cognitive ability, and risk preferences has also been a topic of interest in the literature. Dohmen et al. (2010) finds that individuals with lower cognitive ability tend to be more impatient. However, patience is not only an inherent characteristic but it can also be influenced by several factors. Galor & Özak (2016) examines the determinants of differences in time preferences across countries using LTO. They find that individuals whose ancestors had higher crop yields tend to have higher LTO and behave accordingly by saving more and smoking less. Figlio et al. (2019) analyses the role of LTO on the educational attainments of immigrant students in Florida and finds that students from countries with LTO attitudes have higher exam scores and fewer absences and disciplinary incidents. Doepke & Zilibotti (2008) shows that patience could be transmitted from parents to children. Similarly, S. Chowdhury et al. (2022) provides empirical evidence that parents, spouses, and community norms influence children's patience levels. Furthermore, Alan & Ertac (2018) conducts a randomised study showing that patience can be learned, suggesting that external factors can shape an individual's level of patience. Therefore, we can expect that the society that we live in can affect the patience levels of individuals and hence their decisions in different areas including investment, education and the labour market.

Le Pargneux & Zeitoun (2023) demonstrates that patience is a determinant of differences in subjective well-being across countries. To mitigate potential endogeneity issues, they employ an instrumental variable approach, using the proportion of Protestants in each country as an instrument, based on previous literature linking Protestantism with patience.[2] They find a causal positive effect of patience on subjective well-being that suggests countries with higher levels of patience tend to have higher levels of life satisfaction and happiness. They argue that patient people may be more likely to make decisions that prioritise long-term benefits over short-term gains which leads to greater life satisfaction in the long run.

Dohmen et al. (2016) and Sunde et al. (2022) both use the patience data from the Global Preference Survey (Falk et al., 2018) to investigate the relationship between patience and comparative development across countries. Dohmen et al. (2016) documents a strong positive correlation between patience levels and per capita income, human and physical capital accumulation, productivity, and institutional quality. The study also explores the empirical relevance of the hypothesis that patience influences national income through accumulation processes. It finds strong correlations between patience and human and physical capital accumulation, investments in productivity, and institutional quality.

Building upon these findings, Sunde et al. (2022) presents a more comprehensive analysis by demonstrating that the patience and development relationship holds not only across countries but also across regions within countries and across individuals within regions.

[2]According to Weber (1930), the Protestant emphasis on hard work, thrift, and delayed gratification led to increased savings, investments, and economic growth. This hypothesis has been widely debated and has inspired numerous studies investigating the relationship between religion, cultural values, and economic outcomes.

Interestingly, Sunde et al. (2022) finds that the patience effect on development gets stronger at higher levels of aggregation which means that patience matters more for differences between countries than for differences between individuals.

Similarly, Hübner & Vannoorenberghe (2015) presents evidence that patience is a significant determinant of long-term income disparities between countries, using the patience data from Wang et al. (2016). To address potential endogeneity bias, they instrument patience using the information on how the languages spoken in the countries require speakers to encode time, namely future-term reference (FTR) from M. K. Chen (2013). Their findings indicate that a one-standard-deviation increase in patience can lead to a 34% to 78% increase in per-capita income which supports the role of patience in economic development.

On the other side, the consequences of impatience are far-reaching. Sutter et al. (2013) demonstrates that impatient children and adolescents are more likely to engage in risky behaviours, such as alcohol and cigarette consumption, have higher body mass indices, are less likely to save money and exhibit poorer conduct at school. Moreover, Cadena & Keys (2015) finds that impatient individuals are more likely to drop out of school even when there is less than a year to graduation which results in lower human capital accumulation, lower earnings, and greater regret.

Blanchard & Fischer (1989) suggests that more patient countries save more and attract more foreign assets. Similarly, Nieminen (2022) supports this view by showing that patient countries tend to have current account surpluses, leading to the accumulation of foreign assets. In football, players and managers can be considered foreign assets, as they can be bought and sold in the sports labour market. As a result, patient countries may be more likely to attract talented foreign players and coaches.

The presence of foreign talent can have a positive impact on team performance. Ingersoll et al. (2017) finds that teams with more heterogeneous players, as measured by linguistic diversity, outperformed less heterogeneous teams in the UEFA Champions League. This suggests that the ability to attract and integrate diverse talent can provide a competitive advantage. Moreover, the quality of coaching plays a crucial role in team success. Frick & Simmons (2008) shows that teams hiring better quality coaches can expect a performance improvement, and Szymanski et al. (2019) further demonstrates that coaches with multicultural backgrounds tend to be more successful in highly competitive global environments in football.

As a result, it can be argued that teams from patient countries may have an advantage in attracting talented foreign players and coaches. As patient countries accumulate foreign assets, they may be more willing and able to invest in high-quality talent from abroad. This, in turn, can lead to improved team performance, as diverse player backgrounds and experienced coaches contribute to success on the field. While players are more commonly transferred than managers, the general principle of patience leading to the accumulation of valuable human capital may apply to both.

Moreover, patience can affect labour markets both directly and indirectly by affecting financial situations (Galor & Özak, 2016; Falk et al., 2018), education (Figlio et al., 2019) and well-being (Le Pargneux & Zeitoun, 2023). Therefore, people may be willing to wait longer times to find optimal jobs for their careers. Also, firms may be patient and provide longer periods for their managers and workers to succeed at work. However, to the best of my knowledge, there is no study examining the relationship between patience and labour market activities.

3.2.2 Football Manager Dismissals

The high turnover rate of managers in professional football has attracted considerable research attention over the years. Studies on this topic have primarily focused on two main areas: the determinants of managerial dismissals and the effects of these dismissals on team performance. Understanding the factors that lead to managerial sackings and the consequences of these decisions is crucial for clubs to balance short-term performance pressures with long-term stability and success.

Determinants of Dismissals

The determinants of managerial dismissals in football have been a topic of growing interest in the literature. A range of factors, from short-term performance fluctuations to the influence of social media, have been identified as contributing to the high turnover rate of managers in the sport. However, the existing literature has largely overlooked the role of socioeconomic factors, such as the patience level of society and the pressure exerted by fans, in shaping the decision-making process in managerial dismissals.

One of the earliest studies to investigate this issue was by Audas et al. (1999), which estimates hazard functions for involuntary and voluntary managerial job termination in English professional soccer from 1972 to 1997. They find that short-term fluctuations in performance strongly influence the involuntary termination hazard, along with the team's current league position relative to its position when the manager took charge and the win ratio over the entire spell. Interestingly, they also find that managerial human capital attributes have a greater influence on voluntary rather than involuntary termination.

Several studies have corroborated and expanded upon these findings. Mourao & Araújo (2024) shows that a higher average cost per point increases the chances of a manager being fired. Similarly, Frick et al. (2010) finds that wealthier teams are more likely to dismiss their managers. Moreover, Barros et al. (2009) finds that while the salary of the managers is insignificant for dismissal, high team wage bills are associated with shorter tenures for managers.

The timing of poor performance also plays a role in managerial dismissals. Elaad et al. (2018) shows that the significance of match results for dismissal gradually decreases as more time passes since the match, suggesting that hope and patience for success run out over time. d'Addona & Kind (2014) provides further evidence for this, showing that the possibility of being fired increased from 1949 to 2008 which is the period their data cover. This increase could potentially be a result of declining patience in general. They also find that older managers have a higher probability of being fired and that recent matches are more important for firing decisions. Additionally, managers of teams in the relegation position face a higher probability of getting fired. Bryson, Buraimo, et al. (2021) finds that head coaches' probabilities of dismissal are significantly lower when the team is performing above expectations, with the effect strongest for recent matches.

The non-random nature of coach dismissals, with underperforming clubs more likely to change managers, poses challenges for causal inference. Narita et al. (2023) uses propensity score analysis to obtain counterfactuals and finds that the probability of managerial change increases when a club has lost the last match, performed poorly in the previous four matches, suffered negative surprising results during the season or faced a threat of relegation.

In addition to these factors, sociocultural elements have also been examined. Foroughi et al. (2018) explores the influence of a socioeconomic factor on the turnover frequency of head coaches of national football teams using Hofstede's Long-Term Orientation (LTO) index. They find that countries with a pragmatic culture (high LTO) tend to have less frequent coach changes compared to those with a normative culture (low LTO), although the effect was only significant at the 10% level and the study did not control for coach-specific factors or team performance.

Finally, the influence of social media has emerged as a novel factor in managerial dismissals. Attié et al. (2023) uses machine learning models to show that social media pressure from fans on Twitter is associated with managerial sacking, implying that fan pressure, which could be influenced by patience levels of societies, on social media could determine the tenure duration of football managers.

In summary, the literature has identified a complex web of factors contributing to managerial dismissals in football, from short-term performance and timing of results to the socioeconomic factors and the growing influence of social media.

Effects of Dismissals

The effects of managerial dismissals in football have been the subject of considerable research, with studies yielding mixed results. While some suggest that dismissals can lead to improved performance, others find no significant effect or even negative consequences.

Early work by Audas et al. (2002), using over a quarter-century of match-level data from English football, finds that teams that changed their manager within-season underperformed over the following 3 months on average. They also note that managerial change increased the variance of the non-systematic component of performance in the short term which suggests that the high incidence of within-season managerial change in English football may be a consequence of team owners gambling on this increased variance to avoid relegation.

Several studies have found positive effects of managerial dismissals. Madum (2016) observes positive results after manager dismissals in Denmark, attributing this to the limited number of firings per year, possibly due to the high level of patience in the country, suggesting that changes are only made when necessary. Kattuman et al. (2019) qualitatively shows that new managers can provide increased support to players, restoring lost team motivation. Soebbing et al. (2015), using betting odds as a measure of performance expectations in German football, finds significant positive time-lagged effects on performance expectations 8 weeks after a manager change which indicates that a substantial amount of time is required for increased performance.

However, other studies have found no effect or negative consequences. For example, Hughes et al. (2010) suggests that dismissals may only postpone failure, with no long-term changes. Besters et al. (2016) finds no effect of in-season head coach changes on performance in the English Premier League, supporting the scapegoating theory. van Ours & van Tuijl (2016) also reports a lack of improvement in team performance following managerial dismissals, noting that any apparent improvements were mirrored in the control group which means teams that retained their managers also experienced similar performance improvements. Similarly, de Paola & Scoppa (2012) finds no effect of dismissal when employing the nearest neighbour matching method for analysis.

Some other studies find that the impact of dismissals may be short-lived. de Dios Tena & Forrest (2007) finds that the positive effect of a new manager is limited to home matches and only lasts for 1-2 matches. Radzimiński et al. (2022) and Ponce-Bordón et al. (2023) also find performance increases with a new manager, but only for a short period, possibly due to players not putting in maximum effort under the former manager.

Dismissals can also have unintended consequences. Guerrero-Calderón et al. (2021) notes a decrease in performance in training and matches after dismissals, possibly due to a lack of stability and patience for the required time to implement new strategies. Furthermore, Ekstrand et al. (2023) finds that hamstring injuries increase with new managers, possibly due to increased player effort to impress the new manager.

On the other side, the characteristics of the new manager can also influence post-dismissal performance. Narita et al. (2023) finds that a new manager with a stronger association with the club, such as having finished their playing career there, can positively influence post-succession performance while replacing a manager with a former vice coach tends to have a detrimental effect.

Finally, Flepp & Franck (2021) differentiates between wise dismissals (following actual poor performance) and unwise dismissals (following seemingly poor performance due to bad luck). They find that wise dismissals increased subsequent performance compared to non-dismissals with similarly poor performance, while unwise dismissals did not improve performance compared to a control group with similar bad luck.

The literature reveals that the scapegoating theory provides a valuable framework for understanding football manager dismissals. It suggests that managers are often fired as a way for clubs to deflect blame and appease stakeholders, even if the manager's performance is not the primary cause of the team's poor results. This theory is particularly relevant in societies with low levels of patience, where the pressure for immediate success is high, and the tolerance for underperformance is low. In such contexts, clubs may be more likely to dismiss managers as a symbolic act to demonstrate that they are taking action even if the manager is not solely responsible for the team's performance.

While Foroughi et al. (2018) explores the impact of long-term orientation, a cultural dimension related to patience, on the turnover frequency of national team head coaches, there remains a gap in the current research regarding the influence of patience levels on the tenure of club managers across different countries. Moreover, their study uses the long-term orientation of Hofstede (2001) and employs OLS and negative binomial which prevent them from making causal inferences as there are potential endogeneity issues as long-term orientation is a snapshot variable while head coach data is time-varying. Also, they do not control for success and other characteristics of head coaches. It can be argued that the success of a coach is a determinant of dismissals (Bryson, Buraimo, et al., 2021; Elaad et al., 2018; d'Addona & Kind, 2014). Finally, their findings are significant under 10%. This study aims to fill the gap by examining how the patience levels of various countries affect the job security and decision-making of football managers at the club level. By using a direct measure of patience from Falk et al. (2018), this research provides a more precise analysis of how this specific cultural trait impacts managerial tenure. By investigating this relationship, the study seeks to understand better how socioeconomic factors, particularly patience, shape managerial stability and long-term decision-making in football. Ultimately, the findings of this study may provide insights into the potential consequences of impatience on managerial performance and organisational success, both in football and in other high-pressure environments.

3.3 Data and Methodology

3.3.1 Data

The primary measure of patience has been obtained from the Global Preferences Survey (GPS) conducted by Falk et al. (2018), which covers 76 countries and represents 90% of the world population. This comprehensive survey provides a reliable and standardised measure of patience across different cultures and economies. The patience measure in GPS is derived from a combination of qualitative and quantitative assessments. In the qualitative assessment, respondents are asked to rate their willingness to wait on an 11-point Likert scale, with 0 indicating "completely unwilling to wait" and 10 indicating 'very willing to wait". The quantitative assessment employs a staircase method, where respondents make five consecutive hypothetical choices between receiving a fixed payment today or a larger payment in 12 months. The monetary amounts used in these hypothetical choices are adjusted based on each country's economic conditions, such as per capita income and purchasing power parity, to ensure that the rewards are meaningful and comparable across different countries. This approach allows for a more accurate assessment of individuals' time preferences, accounting for the relative value of money in each country. The combination of these two assessments provides a robust and validated measure of patience at the individual level, which is then aggregated to the country level. For robustness checks, the patience measure of Wang et al. (2016) and the long-term orientation (LTO) dimension of Hofstede (2001) are also employed.

The data on the careers of football managers and match statistics have been obtained from Transfermarkt, a widely used platform for football-related studies. The dataset includes detailed information on managerial tenures, team performance, trophies won, and other relevant variables that are essential for the analysis. Betting odds, used to capture team qualities and calculate expected performance using the cumulative surprise measure of van Ours & van Tuijl (2016), have been obtained from OddsPortal. The cumulative surprise measure is calculated as the difference between the actual number of points earned by a team and the expected number of points based on betting odds. This measure accumulates over the course of a manager's tenure, providing a way to assess the manager's performance relative to expectations. Lastly, economic indicators of countries such as GDP and inflation have been obtained from the World Bank.

This study employs two datasets to investigate the relationship between country-level patience and the tenure of football managers. The first dataset is a career-level dataset that contains one observation for each manager-team spell, providing information on the manager's performance, total days and matches with the team, and other relevant variables. This dataset offers an overview of a manager's career progression and allows for the examination of factors that influence the length of their tenure at each club. Table 3.1 presents the descriptive statistics of the career-level data below. The dataset contains $60,982$ observations from 47 countries which covers 72% of the world population.

The second dataset is a match-level dataset, where each observation represents a single match played by a team. This dataset includes information such as the opponent, betting odds, match score, and whether the manager was dismissed after the match. The match-level data enables a more granular analysis of the factors that contribute to a manager's dismissal, taking into account the performance and circumstances surrounding individual matches. The match-level dataset has 363,473 observations from 35 countries from year 1989 to 2024. The career-level dataset is used for the OLS and IV analyses, which examine the overall impact of patience on tenure, while the match-level dataset is employed in the Accelerated Failure Time (AFT) model, which investigates the factors that influence the timing of manager dismissals. An advantage of using the match-level dataset is controlling performance relative to the expectations by using betting odds for each match, following the methodology outlined in van Ours & van Tuijl (2016) and above.

Table 3.1 *Descriptive Statistics of Career-Level Data*

Statistic	N	Mean	St. Dev.	Min	Max
Manager-Team Related Statistics					
Age at Appointment	59,048	46.443	8.124	17	81
Foreign Manager	60,932	0.237	0.425	0	1
Days in Post	60,887	437.453	905.745	0	17,764
Total Matches	60,932	24.053	41.146	0	1,490
Total Wins	60,932	9.560	19.256	0	895
Total Draws	60,932	6.033	10.472	0	323
Total Losses	60,932	8.460	13.248	0	272
Points per Match	60,932	0.920	0.733	0	3
Country-Related Statistics					
Patience (GPS)	52,509	0.092	0.365	−0.431	1.071
Patience (Wang)	37,875	0.611	0.164	0.080	0.890
Long Term Orientation	59,520	43.389	16.232	1	100
Share of Protestants	52,509	0.084	0.153	0	0.629
Future-Term Reference (FTR)	44,169	0.744	0.422	0	1
Selected Controls					
Temperature	51,387	12.959	6.770	−7.929	27.368
Precipitation	51,387	72.124	40.252	2.911	220.371
Cognitive Skills	48,439	4.569	0.508	3.089	5.338
Satisfaction with Life (Rank)	58,836	73.213	45.157	1	174
Inflation	60,766	15.550	36.023	0.700	360
GDP per capita (in nominal USD)	60,517	27,980.840	21,730.940	504.038	93,446.430
Years Schooling	52,329	9.727	2.101	4.095	13.424
Democracy Index	52,509	8.429	2.358	0	10
Life Expectancy	52,509	77.724	4.774	52.105	83.096
Colonized	52,509	0.358	0.479	0	1
Unemployment	49,926	10.111	6.376	0.510	24.790
Freedom Index	49,926	0.714	0.166	0.373	0.956
Social Support	49,926	0.868	0.079	0.511	0.948

Note: Data sources: Falk et al. (2018), Wang et al. (2016), Hofstede (2001), TransferMarkt & World Bank. Descriptive statistics of additional controls can be found in match-level data in Table C.1 in the Appendix.

3.3.2 Methodology

I begin the analysis by estimating the relationship between country-level patience and manager tenure using a simple Ordinary Least Squares (OLS) regression and career-level dataset. The OLS regression provides a baseline estimate of the association between patience and manager tenure, without accounting for potential endogeneity issues. The regression equation is as follows:

$$\text{Tenure}_{ijt} = \beta_0 + \beta_1 \times \text{Patience}_j + \beta_2 \times \mathbf{X}_{ijt} + \alpha_{\text{i}} + \varepsilon_{ijt} \tag{3.1}$$

Where:

- Tenure_{ijt} represents the tenure of manager i in country j at time t, measured in total matches as the performance of managers evaluated for their match performance.
- Patience$j$ is the measure of patience for country j, obtained from the Global Preferences Survey.
- $\mathbf{X}ijt$ is a vector of control variables that includes:
  - Manager characteristics (e.g., age, foreign dummy).
  - Team-related controls (e.g., average points and trophies won).
  - Country-level controls (e.g., GDP per capita, unemployment rate).
- β_0 is the constant.
- β_1 is the coefficient of interest, representing the association between country-level patience and the tenure of football managers.
- β_2 is a vector of coefficients for the control variables.
- α_{i} is manager fixed effects capturing unobserved heterogeneity for managers.
- ε_{ijt} is the error term.

The methodological approach in this study addresses challenges inherent in analysing the relationship between societal patience and managerial tenure. A simple OLS regression would be inadequate as the patience measure used in this study is a snapshot from the GPS while the manager career data spans multiple years. This mismatch between the static nature of the patience measure and the dynamic nature of the manager tenure data may introduce endogeneity concerns as the patience level of a society could potentially change over time. This issue makes causal inference challenging with OLS. To overcome this, I primarily employ an instrumental variable (IV) approach which helps to establish causality by using exogenous instruments. The IV model serves as the main specification which uses career-level data to examine the effect of patience on managerial tenure.

Instrumental Variable Approach

To establish a causal link between country-level patience and manager tenure, the study employs an instrumental variable (IV) approach as there are potential endogeneity problems that may arise due to omitted variable bias or measurement error. For instance, unobserved factors such as a country's political or economic instability or cultural attitudes may influence both the patience level of the society and the tenure of managers. Moreover, the patience measure used in this study is a snapshot from the Global Preferences Survey, which was conducted at a specific point in time. In contrast, the manager career data is panel data that spans multiple years. This mismatch between the static nature of the patience measure and the dynamic nature of the manager tenure data may introduce additional endogeneity concerns, as the patience level of a society could potentially evolve over time which is not captured by the measure used in this study. To address these potential endogeneity issues, I employ an instrumental variable approach using two different instrumentals.

The first IV is the future-term reference (FTR) of languages which captures the linguistic encoding of time in the languages spoken in each country and it is a linguistic feature that indicates whether a language requires future events to be grammatically marked. In languages with strong FTR, such as English, speakers must use the future tense to discuss future events (e.g., "I will go to the football match tomorrow"). In contrast, languages with weak FTR, like German, do not require explicit future tense marking (e.g., "Ich gehe morgen zum Fußballspiel" - literally, "I go to the football match tomorrow"). These differences in language structure may lead to varying perceptions of time across language groups. For the first time, M. K. Chen (2013) discovers that speakers of languages with weak FTR exhibit more future-oriented behaviours, such as increased savings, greater retirement wealth, reduced smoking, safer sexual practices, and lower obesity rates, compared to those who speak languages with strong FTR. These associations were observed both across and within countries, even when comparing demographically similar native households. Building on these findings, Hübner & Vannoorenberghe (2015) employs strong FTR as an instrumental variable for patience in their study. Similarly, I use the population-weighted average of the strong FTR as an instrument for patience as there could be more than one spoken language in a country.

The second IV is the share of Protestants in a country, based on the historical association between Protestantism and a strong work ethic, thrift, and delayed gratification (Weber, 1930) The Protestant Reformation emphasized the importance of hard work, frugality, and investing in the future, which is closely related to the concept of patience. The share of Protestants in a country previously used as an instrument for patience by Dohmen et al. (2016) and Le Pargneux & Zeitoun (2023).

The strong FTR and the share of Protestants are expected to be correlated with patience but uncorrelated with the error term in the second-stage regression to provide exogenous variation in patience levels. I expect patience to be negatively correlated with the strong FTR and positively correlated with the share of Protestants. In the instrumental variable analysis, I primarily use strong FTR as the main instrument for patience, while the share of Protestants in a country is used as an alternative instrument to demonstrate the robustness of the findings.

The IV estimation consists of two stages: the first stage, where patience is regressed on the instrumental variables and control variables, and the second stage, where manager tenure is regressed on the predicted values of patience from the first stage and control variables.

The first-stage regression equation for strong FTR as IV is as follows:[3]

$$\text{Patience}_j = \gamma_0 + \gamma_1 \times \text{FTR}_j + \gamma_2 \times \mathbf{X}_{ijt} + \alpha_i + u_{ijt} \tag{3.2}$$

Where:

- Patience_j is the level of patience in country j.
- FTR_j is the population-weighted average of strong FTR of spoken languages in country j.
- γ_0 is the constant.
- γ_1 is the coefficient of interest in the first-stage regression which shows the association between the instrumental variable and patience.
- γ_2 is the vector of coefficients for the control variables that include geographical, colonial, linguistic, religious, institutional and socioeconomic variables.
- $\mathbf{X}_{ijt}$ is a vector of control variables.
- α_i is manager fixed effects.
- u_{ijt} is the error term in the first-stage regressions.

Second-stage regression equation:

$$\text{Tenure}_{ijt} = \beta_0 + \beta_1 \times \widehat{\text{Patience}}_j + \beta_2 \times \mathbf{X}_{ijt} + \alpha_i + \varepsilon_{ijt} \tag{3.3}$$

Where:

- Tenure_{ijt} is the total tenure duration of manager (in matches) i at team t in country j.

[3] The first-stage regression equation for the share of Protestants as IV replaces FTR in the equation with the share of Protestants in country j.

- $\widehat{\text{Patience}}_j$ is the predicted value of patience for country j from the first-stage regression.
- β_0 is the constant.
- β_1 is the coefficient of interest, capturing the causal effect of patience on manager tenure.
- β_2 is a vector of coefficients for the control variables.
- $\mathbf{X}_{ijt}$ is a vector of control variables.
- α_i is manager fixed effects.
- ε_{ijt} is the error term.

As a robustness check, I also conducted a survival analysis, an accelerated failure time (AFT) model. The AFT model accounts for right-censoring in the data which is a crucial factor given that some managers in the sample have not been dismissed by the end of the observation period. Specifically, 3.09% of the observations represent managers who were still in their positions at the end of the observation period. Although this proportion is not large, I use the AFT model to check the robustness of the results. Importantly, the AFT model uses match-level data, providing a different perspective on the relationship between patience and managerial tenures. The use of these two distinct datasets necessitates careful interpretation of the results. A key strength of the identification strategy is the managers who switch countries during their careers which provides within-manager variation across different patience levels in countries. This cross-country movement provides an opportunity to isolate the effect of patience on tenure while controlling for individual manager characteristics.

Survival Analysis

In addition to the IV analysis, I conduct a survival analysis using the accelerated failure time (AFT) model. The AFT model is a suitable choice for this study, as it allows for the examination of the factors influencing the duration of managerial tenures while accounting for censored observations (i.e., managers who are not dismissed yet at the end of the study period). The AFT model estimates the effect of country-level patience on the time to managerial dismissal, controlling for variables such as team performance, manager characteristics, and country-specific factors using the match-level dataset.

The choice of the AFT model over the more commonly used Cox proportional hazards model is based on the assumption of proportional hazards. The Cox model assumes that the hazard ratio between groups remains constant over time. However, in the case of football managers, it is possible that certain factors could accelerate or decelerate the survival rate of managers which violates the proportional hazards assumption (Bryson, Buraimo, et al., 2021). For example, a highly successful manager may experience a slower rate of dismissal risk over time compared to a less successful one. The AFT model, on the other hand, does not rely on the proportional hazards assumption and allows for the effect of covariates to vary over time, making it more flexible and appropriate for this study (Faruk, 2018). Based on the Akaike Information Criterion (AIC), I chose the Weibull distribution for the AFT model as it fits the data best. The regression equation for the AFT model is given below:

$$\ln(T_{ijt}) = \beta_0 + \beta_1 \times \text{Patience}_j + \beta_2 \times \mathbf{X}_{ijt} + \alpha_i + \varepsilon_{ijt} \tag{3.4}$$

Where:

- T_{ijt} is the survival time (tenure) of manager i in country j at time t.
- Patience_j is the measure of patience for country j.
- $\mathbf{X}_{ijt}$ is a vector of control variables that includes manager, team and country-related controls.
- β_0 is the constant.

- β_1 is the coefficient of interest, capturing the effect of country-level patience on the logarithm of manager tenure.
- β_2 is a vector of coefficients for the control variables.
- α_i is the manager fixed effects.
- ε_{ijt} is the error term.

3.4 Results

3.4.1 Ordinary Least Squares (OLS)

The results of OLS are given below in Table 3.2. The main variable of interest, patience, is positively and statistically significantly associated with the total number of matches coached by a manager across all model specifications. In the most comprehensive model (column 4), which includes manager-level controls, country-level controls, and a fixed effect for managers, a one-unit increase in the patience level of a country is associated with managers coaching approximately 13.52 more matches on average.

The control variables, such as points per match and age at appointment, have the expected signs and are statistically significant. The inclusion of country-level controls such as GDP per capita, unemployment and life expectancy which could affect the patience level of countries in the model (4) does not substantially change the magnitude or significance of the coefficient of patience. Overall, these results provide evidence for a positive relationship between patience and the length of football manager tenures, but further analysis is needed to establish causality.

Table 3.2 *The Effect of Patience on Manager Tenure: OLS Results*

Dependent Variable:	Total Matches Coached by Manager			
Model:	(1)	(2)	(3)	(4)
Patience	22.16***	15.76***	14.65***	13.52***
	(3.475)	(2.795)	(2.678)	(2.884)
Points per Match		18.77***	18.73***	18.72***
		(1.918)	(1.913)	(1.915)
Foreign Manager		-5.642***	-3.267	-3.463*
		(1.831)	(1.949)	(1.909)
Age at Appointment		-0.6732***	-0.6570***	-0.6504***
		(0.1711)	(0.1710)	(0.1732)
Country-Specific Controls	No	No	No	Yes
Fixed-effects				
Manager	Yes	Yes	Yes	Yes
Fit statistics				
Observations	49,926	48,643	48,259	48,168
R^2	0.481	0.535	0.536	0.536

Data sources: Falk et al. (2018), TransferMarkt, World Bank. Career-level data. Country-specific controls include GDP per capita, inflation, unemployment, democracy index, and life expectancy among others. A full list of controls can be found in Table C.1 in the Appendix. Standard errors clustered at the country level are in parentheses. The decrease in observations from Model (1) to Model (4) is due to missing data on the age of managers.
* p<0.1; ** p<0.05; *** p<0.01.

The OLS regression provides a simple and intuitive way to estimate the relationship between patience and manager tenure. However, the OLS estimates may be biased due to the endogeneity problem. To address potential endogeneity problems, I employ an instrumental variable approach, which will be discussed in the following section.

3.4.2 Instrumental Variable (IV)

The instrumental variable (IV) analysis was employed to address the endogeneity concerns inherent in the OLS regression, using exogenous variation in patience levels provided by two different instruments. This approach ensures a more reliable estimation of the causal effect of patience on the tenure of football managers. The instruments used in this analysis are the future-time reference (FTR) of languages and the share of

Protestants in a country. The choice of these instruments is grounded in their strong theoretical rationale and their plausible exogeneity with respect to football managers' tenure, apart from their impact on patience levels. Their validity is further corroborated by a correlation analysis, graphically illustrated in Figures C.1 and C.2 in the Appendix, demonstrating that the FTR and the share of Protestants are unlikely to directly influence manager tenure apart from through their effect on patience.

Table 3.3 presents the results of the instrumental variable (IV) analysis using career-level data and both instruments: the future-time reference (FTR) of languages and the share of Protestants in a country. Both instruments provide consistent and statistically significant estimates of the causal effect of patience on the tenure of football managers, as measured by the number of matches coached during the tenure.

When using the FTR as an instrument, a one-unit increase in the patience level of a country leads to managers coaching approximately 6 more matches on average, holding other factors constant. Similarly, when using the share of Protestants as an instrument, a one-unit increase in patience results in managers coaching about 7 more matches. These findings confirm that there is a positive relationship between patience and manager tenure as observed in the OLS regression.

As expected, points per match, a measure of manager performance, has a strong positive effect on tenure which highlights the importance of on-field success in determining manager job security.

Table 3.3 *The Effect of Patience on Manager Tenure: IV Results*

	Dependent variable:	
	Number of Matches	
Instrument:	(FTR)	(Share of Protestants)
Patience	6.218***	7.358***
	(1.206)	(1.000)
Points per Match	24.466***	22.612***
	(0.301)	(0.248)
Foreign Manager	−0.037	0.083
	(0.619)	(0.471)
Age at Appointment	−0.276***	−0.222***
	(0.027)	(0.022)
Country-Specific Controls	Yes	Yes
First-Stage F-Statistic	104442.7	144939.2
Fixed-effects		
Manager	Yes	Yes
Observations	38,680	48,168
R^2	0.216	0.221

Note: Data sources: Falk et al. (2018), Le Pargneux & Zeitoun (2023), TransferMarkt, World Bank. Career-Level data. Country-Related controls include GDP per capita, inflation, unemployment, democracy index, and life expectancy among others. A full list of controls can be found in Table C.1 in the Appendix. Standard errors clustered at the country level are in parentheses. The difference in observations between the two models is due to data availability for the instrumental variables. * p<0.1; ** p<0.05; *** p<0.01.

Similar to the OLS, being foreign is not associated with the tenure times of managers when the controls are added in both IV results. The age at appointment has a negative effect on tenure, indicating that managers appointed at an older age tend to have shorter tenures. This may be because clubs hiring young managers allow more time for them to adapt to challenges and implement new strategies.

For robustness, I replicated the analysis using the match-level dataset and controls and these findings are robust at match-level data where being dismissed after the match is the dependent variable. As being dismissed is a binary variable, after the first-stage, I use logistic regression with two different approaches: Two-Stage Predictor Substitution

(2SPS) and Two-Stage Residual Inclusion (2SRI) (Terza et al., 2008). The results of this robustness test are presented in Table C.2 in the Appendix.

3.4.3 The Accelerated Failure Time (AFT) Model

Besides instrumental variables, an AFT model is also employed to examine the factors affecting the time to dismissal of football managers, with a particular focus on the role of patience. Table 3.4 presents the AFT results for two different specifications: one is for all dismissals and another for in-season dismissals only. In-season dismissals are more likely to be direct firings, reflecting an immediate response to poor performance, and pressure from fans and media. These dismissals are often reactive and may be influenced by the impatience of the club's stakeholders, who demand action to address perceived problems.

On the other side, end-of-season dismissals may be driven by different factors, such as the expiration of contracts or a more comprehensive evaluation of the manager's performance over the entire season. Clubs may be more willing to allow managers to complete the season and assess their performance at the end. End-of-season dismissals could be biased and may not show the real effect of patience as teams may avoid dismissing managers if only a few matches are left until the end of the season. Because of that reason, in-season dismissals analysis could be more reliable.

Table 3.4 *The Effect of Patience on Manager Tenure: AFT Results*

	Time to Dismissal	
	(All Dismissals)	(In-Season Dismissals)
Patience	2.393***	2.839***
	(0.070)	(0.104)
Foreign Manager	−0.069***	−0.078**
	(0.022)	(0.031)
Home	0.014	0.023
	(0.015)	(0.022)
League Tier	−0.060***	−0.074**
	(0.021)	(0.029)
Age at Appointment	0.009***	0.014*
	(0.003)	(0.007)
FIFA Rank	−0.022	−0.021
	(0.028)	(0.027)
Cumulative Surprise	0.029***	0.036***
	(0.006)	(0.008)
Team Promoted	−1.727***	1.717*
	(0.490)	(0.940)
Won Domestic Super Cup	2.463***	2.368***
	(0.352)	(0.400)
Won Domestic Cup	2.367***	3.412***
	(0.455)	(1.018)
Won Continental Cup 1	2.976***	15.775***
	(1.06)	(0.940)
Won Continental Super Cup	0.735**	0.503
	(0.303)	(1.091)
Won Continental Cup 2	4.434***	21.900***
	(0.237)	(0.961)
Won Domestic League	−0.299*	−0.519***
	(0.170)	(0.227)
Country-Specific Controls	Yes	Yes
Observations	270,912	270,912
Log Likelihood	−202,535.800	−106,626.700
χ^2 (df = 22)	12,394.800***	7,134.662***

Note: Data sources: Falk et al. (2018), TransferMarkt, World Bank. Match-Level data. Country-Related controls include GDP per capita, inflation, unemployment, democracy index, and life expectancy among others. A full list of controls can be found in Table C.1 in the Appendix. Robust standard errors are in parentheses.

* p<0.1; ** p<0.05; *** p<0.01.

In both specifications, patience has a strong positive effect on the time to dismissal. A one-unit increase in patience is associated with $e^{2.393} \approx 11$ more matches to dismissal when considering all dismissals and $e^{2.839} \approx 17$ increase in number of matches when focusing on in-season dismissals. This difference could stem from the reason mentioned above.

Several other factors also affect the time to dismissal. Similar to the IV analysis, foreign managers tend to have shorter times to dismissal compared to domestic managers, with the effect being more pronounced for in-season dismissals. Managers coaching teams in higher tiers (leagues attracting more interest) face a higher risk of dismissal as fan and media pressure might be higher there.

Performance-related variables have the expected effects on the time to dismissal. A higher cumulative surprise, which captures the team's performance relative to expectations during the manager's spell, is associated with a longer time to dismissal. Winning domestic cups, continental cups, and super cups in the previous year also increases the time to dismissal, highlighting the importance of tangible successes in securing a manager's job. However, winning the domestic league, similar to promoting from a lower division in the previous season, has a negative effect on the time to dismissal, possibly due to increased expectations of fans and team owners from the previous year.

In conclusion, the IV and AFT analyses provide compelling evidence for the causal effect of patience on the tenure and time to dismissal of football managers. The results are robust to different measures of patience (Wang et al., 2016; Hofstede, 2001; Falk et al., 2018). Robustness checks with different measures of patience are presented in Tables C.3 and C.4 in the Appendix for IV and AFT, respectively. The findings highlight the importance of patience in shaping the job security and decision-making process of leaders in sports and potentially in other domains. The generalisability of these results to the corporate world suggests that patience may play a crucial role in fostering long-term value creation and reducing short-termism in managerial decision-making.

3.5 Discussion

Considering the average tenure of managers is around 24 matches, the impact of patience on managerial job security is substantial. A one standard deviation increase in patience (0.365 units on a scale from -0.431 to 1.071) leads to managers coaching between 2.19 and 6.21 more matches, depending on the model and data selected. This represents a 9.1% to 25.9% increase in job security relative to the average tenure. These findings suggest that a society's patience level has a meaningful influence on the stability and longevity of managerial careers in football.

This extra time can provide managers and players with crucial opportunities to achieve success in the long run. It allows managers to work on the training ground, try new tactics, and know the strengths and weaknesses of their players. Then, they can understand what works best for the team and make changes accordingly. Moreover, they can transfer new players according to their needs. On the other side, players who have more time with a manager can build a better relationship and understanding. They can learn how the manager wants them to play and what their role is in the team. This specialisation may lead to better performances on the pitch and more success for the club in the long run. Without this extra time as in impatient societies, managers might not be able to fully show what they can do and feel a need for an unnecessary risk. Also, players might not have the chance to train and improve under a consistent coaching style. Moreover, if a player knows that the manager could be fired soon, they might be less motivated to listen to the manager and play according to his tactics. This uncertainty can lead to a lack of commitment and effort from the players, which can decrease the team's performance on the pitch. As shown by Ekstrand et al. (2023), instability in coaching may lead to more injuries which is a temporary human capital loss that hurts performance.

In line with the literature and expectations, a successful manager tends to have longer tenures. On the other side, there appears to be a negative correlation between the age of the managers and the length of their tenure. Given that the age of a manager is not a random attribute, it is prudent to refrain from inferring a causal relationship from this correlation. It could be suggested that clubs hiring younger managers might be adopting a long-term approach which allows for greater tolerance of short-term performance variances.

These findings also show the prevalence of the scapegoating theory especially in impatient societies, where managers are often dismissed as a symbolic act to appease fans and media. This tendency can lead to a vicious cycle of high managerial turnover, job insecurity, and suboptimal decision-making, as managers may prioritise short-term results over long-term ones.

Moreover, the findings provide empirical support for the theoretical framework proposed by Becker & Murphy (2000) on the impact of the social environment on human behaviour. The study's results which show that managers in more patient societies have longer tenures, reflect how societal norms can influence organisational decision-making. This is also evident in the context of the scapegoating theory in football, as discussed by Grusky (1963). The lower managerial turnover in patient societies suggests a reduced tendency towards scapegoating which aligns with the findings of Foroughi et al. (2018) which finds that national team head coaches stay longer in their positions in long-term oriented countries.

The impact of patience on managerial tenure observed in this study is also consistent with studies examining the effects of patience on other economic outcomes. For example, Falk et al. (2018) finds that patience is positively correlated with savings behaviour and educational attainment across countries. Similarly, Figlio et al. (2019) demonstrated that students from cultures with higher long-term orientation perform better academically. In the area of financial decision-making, Wang et al. (2016) shows that patience is associated with higher credit ratings and better environmental protection at the country level. These findings, along with the results of this study on managerial tenure, show the effect of patience on labour market activities. The findings across these different areas suggest that patience is a fundamental cultural trait that shapes behaviour positively in economic activities ranging from individual career trajectories to broader societal outcomes.

3.6 Concluding Remarks

This study focused on the relationship between the patience levels of societies and the tenure of football managers to provide insights into how cultural and socioeconomic factors shape employment practices in high-pressure roles. By using data from football, a sport that reflects societal and cultural norms, the study has shown that patience is a significant determinant of managerial tenure.

The case of Sir Alex Ferguson at Manchester United serves as a prime example of how a society's patience can contribute to the long-term success of a manager. Ferguson's tenure was not only a result of his exceptional skills but also a reflection of the patience and support he received from fans and club owners, especially during the challenging early years. This study has demonstrated that such patience is not universal, and in less patient societies, like Turkey, managers may face a higher risk of dismissal due to the pressure for immediate success.

The analysis has also shown the relevance of the scapegoating theory in manager dismissals, particularly in impatient societies. When faced with poor results, clubs in such societies may resort to firing managers as a symbolic act to appease stakeholders, even if the underlying issues are not directly attributable to the manager. This behaviour can lead to a higher turnover rate for managers and create an environment of job insecurity, which can negatively impact performance and commitment in the workplace.

By employing an instrumental variable approach and survival analysis, this study has established a causal link between a country's patience level and the tenure of football managers. The findings suggest that patience is a crucial factor in allowing managers the time and support needed to implement long-term strategies and build teams accordingly.

The findings of the study are robust to different measures of patience (patience measures of Falk et al. (2018) and Wang et al. (2016) and long-term orientation dimension of Hofstede (2001)).

The findings on the causal relationship between societal patience and football manager tenure may be also applicable to other professional areas. This study demonstrates that higher levels of societal patience directly lead to longer tenures for football managers. Such a causal link likely extends to other high-pressure roles, particularly where performance outcomes are not immediately evident such as CEOs and politicians. In more patient societies, organisations may allow more time for long-term strategies to yield results, reducing the frequency of dismissals. By using data from football, where tenure and performance are easily measured, this study shows how cultural factors directly influence employment stability and decision-making in different professional settings.

Limitations

While this study provides evidence for the causal relationship between patience and manager tenure, it is important to acknowledge its limitations. One limitation is that the patience data used in this study is a snapshot from the Global Preferences Survey, which was conducted at a specific point in time. As societies and cultures evolve, patience levels may change over time, which could affect the long-term applicability of the findings. Furthermore, while the Global Preferences Survey provides a reliable and standardised measure of patience, it is based on survey responses and may not perfectly capture the true level of patience in a society. Despite the use of instrumental variables, control variables, and robustness checks with other measures of patience, there may still be unobserved factors that influence both patience levels and manager tenure, which could not be fully accounted for in this study.

Additionally, the study does not differentiate between managers who voluntarily resign and those who are fired by their clubs, which could provide a more nuanced understanding of how patience influences managerial turnover (van Ours & van Tuijl, 2016). Lastly, although football managers share similarities with CEOs and other executives in other industries, the findings may not be directly generalisable to other contexts without further research (Angrist & Pischke, 2009).

Future Work

This study opens up several avenues for future research. One promising direction is to differentiate between manager resignations and firings, exploring how patience influences each type of turnover separately. This could provide a more comprehensive understanding of the dynamics between patience and managerial turnover.

Another potential avenue for future research is to investigate how changes in societal patience over time may influence the tenure of managers. While this study relies on patience data from a specific point in time, it is possible that patience levels within a society may evolve over time, which could, in turn, affect the job security of managers. For instance, some, including himself, argue that if Sir Alex Ferguson were to manage Manchester United in today's climate, he might face a higher risk of dismissal due to a potential difference in societal patience (Elberse, 2013). A longitudinal study that captures the dynamics of patience over an extended period could shed light on the relationship between changing patience levels and managerial tenure. Such research would require repeated measurements of patience across multiple time points, as well as a corresponding dataset of managerial tenures spanning the same period.

Finally, extending the analysis to other sectors, such as business, politics, or education, could help determine whether the relationship between patience and job tenure holds in different contexts, broadening the findings of this study.

References

Adler, M. (1985). Stardom and Talent. *The American Economic Review*, *75*(1), 208–212.

Akerlof, G. A. (1980). A Theory of Social Custom, of Which Unemployment May be One Consequence. *The Quarterly Journal of Economics*, *94*(4), 749–775.

Akerlof, G. A. (1997). Social Distance and Social Decisions. *Econometrica*, *65*(5), 1005–1027.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics*, *115*(3), 715–753.

Akerlof, G. A., & Kranton, R. E. (2005). Identity and the Economics of Organizations. *Journal of Economic Perspectives*, *19*(1), 9–32.

Akerlof, G. A., & Kranton, R. E. (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton University Press.

Alan, S., & Ertac, S. (2018). Fostering Patience in the Classroom: Results from Randomized Educational Intervention. *Journal of Political Economy*, *126*(5), 1865–1911.

Anand, A., Dalmasso, A., Vessal, S. R., Parameswar, N., Rajasekar, J., & Dhal, M. (2023). The effect of job security, insecurity, and burnout on employee organizational commitment. *Journal of Business Research*, *162*, 113843.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Arcidiacono, P., Kinsler, J., & Price, J. (2017). Productivity Spillovers in Team Production: Evidence from Professional Basketball. *Journal of Labor Economics*, *35*(1), 191–225.

Ariely, D., & Wertenbroch, K. (2002, May). Procrastination, Deadlines, and Performance: Self-Control by Precommitment. *Psychological Science*, *13*(3), 219–224.

Arrow, K. J. (1951). *Social Choice and Individual Values*. Yale University Press.

Attié, M., Pacheco, D., & Oliveira, M. (2023). Getting the Boot? Predicting the Dismissal of Managers in Football. In A. S. Teixeira, F. Botta, J. F. Mendes, R. Menezes, & G. Mangioni (Eds.), *Complex Networks XIV* (pp. 132–140). Springer Nature.

Audas, R., Dobson, S., & Goddard, J. (1999). Organizational Performance and Managerial Turnover. *Managerial and Decision Economics*, *20*(6), 305–318.

Audas, R., Dobson, S., & Goddard, J. (2002). The Impact of Managerial Change on Team Performance in Professional Sports. *Journal of Economics and Business*, *54*(6), 633–650.

Azoulay, P., Graff Zivin, J. S., & Wang, J. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, *125*(2), 549–589.

Bandiera, O., Barankay, I., & Rasul, I. (2010). Social Incentives in the Workplace. *The Review of Economic Studies*, *77*(2), 417–458.

Barros, C. P., Frick, B., & Passos, J. (2009). Coaching for Survival: The Hazards of Head Coach Careers in the German 'Bundesliga'. *Applied Economics*, *41*(25), 3303–3311.

Bass, B. M. (1985). *Leadership and Performance Beyond Expectations*. The Free Press.

Bass, B. M., Avolio, B., Jung, D., & Berson, Y. (2003). Predicting Unit Performance by Assessing Transformational and Transactional Leadership. *Journal of Applied Psychology*, *88*(2), 207–218.

Becker, G. (1957). *The Economics of Discrimination*. University of Chicago Press.

Becker, G., & Murphy, K. (2000). *Social Economics: Market Behavior in a Social Environment*. Harvard University Press.

Berri, D. J., & Schmidt, M. B. (2006). On the Road With the National Basketball Association's Superstar Externality. *Journal of Sports Economics*, *7*(4), 347–358.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, *94*(4), 991–1013.

Besters, L. M., van Ours, J. C., & van Tuijl, M. A. (2016). Effectiveness of In-Season Manager Changes in English Premier League Football. *De Economist*, *164*(3), 335–356.

Bilen, E., & Matros, A. (2023). The Queen's Gambit: Explaining the Superstar Effect Using Evidence from Chess. *Journal of Economic Behavior & Organization*, *215*, 307–324.

Blanchard, O., & Fischer, S. (1989). *Lectures on Macroeconomics*. MIT Press.

Blank, D. B., & Hadley, B. (2021). When CEOs Adapt: An Investigation of Manager Experience, Policy and Performance Following Recessions. *Journal of Corporate Finance*, *71*, 102118.

Bommer, C., Dreher, A., & Perez-Alvarez, M. (2022). Home Bias in Humanitarian Aid: The Role of Regional Favoritism in the Allocation of International Disaster Relief. *Journal of Public Economics*, *208*, 104604.

Bose, P., Feess, E., & Mueller, H. (2022). Favoritism towards High-Status Clubs: Evidence from German Soccer. *The Journal of Law, Economics, and Organization*, *38*(2), 422–478.

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.

Brady, R. R., Insler, M. A., & Rahman, A. S. (2017). Bad Company: Understanding Negative Peer Effects in College Achievement. *European Economic Review*, *98*, 144–168.

Brandes, L., Franck, E., & Nüesch, S. (2008). Local Heroes and Superstars: An Empirical Analysis of Star Attraction in German Soccer. *Journal of Sports Economics*, *9*(3), 266–286.

Brown, J. (2011). Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy*, *119*(5), 982–1013.

Bryson, A., Buraimo, B., Farnell, A., & Simmons, R. (2021). Time To Go? Head Coach Quits and Dismissals in Professional Football. *De Economist*, *169*(1), 81–105.

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.

Bryson, A., Rossi, G., & Simmons, R. (2014). The Migrant Wage Premium in Professional Football: A Superstar Effect? *Kyklos*, *67*(1), 12–28.

Buraimo, B., Forrest, D., & Simmons, R. (2010). The 12th man?: Refereeing Bias in English and German soccer. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *173*(2), 431–449.

Burns, J. M. (1978). *Leadership.* Open Road Media.

Butalia, R., Fransen, K., Coffee, P., Laenens, J., & Boen, F. (2021). Why the Chosen Ones May Not Always Be the Best Leaders: Criteria for Captain Selection as Predictors of Leadership Quality and Acceptance. *Frontiers in Psychology*, *11*.

Butler, R., & Butler, D. (2017). Fergie Time and the Allocation of Additional Time: Evidence from the English Premier League 2009 to 2013. *International Journal of Sport Finance*, *12*(3), 185–203.

Cadena, B. C., & Keys, B. J. (2015). Human Capital and the Lifetime Costs of Impatience. *American Economic Journal: Economic Policy*, *7*(3), 126–153.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, *225*(2), 200–230.

Carozzi, F., & Repetto, L. (2016). Sending the Pork Home: Birth Town Bias in Transfers to Italian Municipalities. *Journal of Public Economics*, *134*, 42–52.

Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual Behavior and Group Membership. *The American Economic Review*, *97*(4), 1340–1352.

Charness, G., & Sutter, M. (2012). Groups Make Better Self-Interested Decisions. *Journal of Economic Perspectives*, *26*(3), 157–176.

Chen, J. S., & Garg, P. (2018). Dancing with the Stars: Benefits of a Star Employee's Temporary Absence for Organizational Performance. *Strategic Management Journal*, *39*(5), 1239–1267.

Chen, M. K. (2013). The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. *American Economic Review*, *103*(2), 690–731.

Chmait, N., Robertson, S., Westerbeek, H., Eime, R., Sellitto, C., & Reid, M. (2020). Tennis superstars: The Relationship Between Star Status and Demand for Tickets. *Sport Management Review*, *23*(2), 330–347.

Chowdhury, S., Sutter, M., & Zimmermann, K. F. (2022). Economic Preferences across Generations and Family Clusters: A Large-Scale Experiment in a Developing Country. *Journal of Political Economy*, *130*(9), 2361–2410.

Chowdhury, S. M., Jewell, S., & Singleton, C. (2024). Can awareness reduce (and reverse) identity-driven bias in judgement? Evidence from international cricket. *Journal of Economic Behavior & Organization*, *226*, 106697.

Cohen-Zada, D., Dayag, I., & Gershoni, N. (2023). Effort Peer Effects in Team Production: Evidence from Professional Football. *Management Science*, *70*(4), 2355–2381.

Cotterill, S. T., & Fransen, K. (2016). Athlete Leadership in Sport Reams: Current Understanding and Future Directions. *International Review of Sport and Exercise Psychology*, *9*(1), 116–133.

Courneya, K. S., & Carron, A. V. (1992). The Home Advantage In Sport Competitions: A Literature Review. *Journal of Sport and Exercise Psychology*, *14*(1), 13–27.

Dagaev, D., Paklina, S. N., Reade, J., & Singleton, C. (2021). *The Iron Curtain and Referee Bias in International Football.* SSRN Electronic Journal.

Dawson, P., & Dobson, S. (2010). The Influence of Social Pressure and Nationality on Individual Decisions: Evidence from the Behaviour of Referees. *Journal of Economic Psychology*, *31*(2), 181–191.

Dawson, P., Dobson, S., Goddard, J., & Wilson, J. (2007). Are Football Referees Really Biased and Inconsistent?: Evidence on the Incidence of Disciplinary Sanction in the English Premier League. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *170*(1), 231–250.

Dawson, P., Massey, P., & Downward, P. (2020). Television Match Officials, Referees, and Home advantage: Evidence from the European Rugby Cup. *Sport Management Review*, *23*(3), 443–454.

de Dios Tena, J., & Forrest, D. (2007). Within-Season Dismissal of Football Coaches: Statistical Analysis of Causes and Consequences. *European Journal of Operational Research*, *181*(1), 362–373.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, *47*(2), 315–372.

DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for Altruism and Social Pressure in Charitable Giving. *The Quarterly Journal of Economics*, *127*(1), 1–56.

de Paola, M., & Scoppa, V. (2012). The Effects of Managerial Turnover: Evidence from Coach Dismissals in Italian Soccer Teams. *Journal of Sports Economics*, *13*(2), 152–168.

Depken, C. A., & Haglund, L. E. (2011). Peer Effects in Team Sports: Empirical Evidence From NCAA Relay Teams. *Journal of Sports Economics*, *12*(1), 3–19.

Deutscher, C., Neuberg, L., & Thiem, S. (2023). Who's Afraid of the GOATs? - Shadow Effects of Tennis Superstars. *Journal of Economic Psychology*, *99*, 102663.

de Chaisemartin, C., & D'Haultfœuille, X. (2023). Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: a Survey. *The Econometrics Journal*, *26*(3), 1–30.

Do, Q.-A., Nguyen, K.-T., & Tran, A. N. (2017). One Mandarin Benefits the Whole Clan: Hometown Favoritism in an Authoritarian Regime. *American Economic Journal: Applied Economics*, *9*(4), 1–29.

Doepke, M., & Zilibotti, F. (2008). Occupational Choice and the Spirit of Capitalism. *The Quarterly Journal of Economics*, *123*(2), 747–793.

Does Football Need a 60-Minute 'Stop-Clock'? (2022). BBC Sport. Retrieved 2023-03-12, from https://www.bbc.com/sport/football/61342349

Dohmen, T. (2008). The Influence of Social Forces: Evidence from the Behavior of Football Referees. *Economic Inquiry*, *46*(3), 411–424.

Dohmen, T., Enke, B., Falk, A., Huffman, D., & Sunde, U. (2016). *Patience and the Wealth of Nations* (Working Paper No. 2015). Human Capital and Economic Opportunity Working Group.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review*, *100*(3), 1238–1260.

Dohmen, T., & Sauermann, J. (2016). Referee Bias. *Journal of Economic Surveys*, *30*(4), 679–695.

Downward, P., & Jones, M. (2007). Effects of Crowd Size on Referee Decisions: Analysis of the FA Cup. *Journal of Sports Sciences*, *25*(14), 1541–1545.

Dreher, A., & Fuchs, A. (2015). Rogue Aid? An Empirical Analysis of China's Aid Allocation. *Canadian Journal of Economics/Revue canadienne d'économique*, *48*(3), 988–1023.

d'Addona, S., & Kind, A. (2014). Forced Manager Turnovers in English Soccer Leagues: A Long-Term Perspective. *Journal of Sports Economics*, *15*(2), 150–179.

Ekstrand, J., Van Zoest, W., & Gauffin, H. (2023). Changes in Head Staff Members in Male Elite-Level Football Teams are Associated with Increased Hamstring Injury Burden for that Season: the UEFA Elite Club Injury Study. *BMJ Open Sport & Exercise Medicine*, *9*(4), 001640.

Elaad, G., Jelnov, A., & Kantor, J. (2018). You Do Not Have to Succeed, Just Do Not Fail: When Do Soccer Coaches Get Fired? *Managerial and Decision Economics*, *39*(3), 269–274.

Elberse, A. (2013). Ferguson's Formula. *Harvard Business Review*.

Emerson, J., & Hill, B. (2018). Peer Effects in Marathon Racing: The role of Pace Setters. *Labour Economics*, *52*, 74–82.

Endrich, M., & Gesche, T. (2020). Home-bias in Referee Decisions: Evidence from "Ghost Matches" During the Covid19-Pandemic. *Economics Letters*, *197*, 109621.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global Evidence on Economic Preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692.

Falk, A., & Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, *24*(1), 39–57.

Faltings, R., Krumer, A., & Lechner, M. (2023). Rot-Jaune-Verde: On Linguistic Bias of Referees in Swiss Soccer. *Kyklos*, *76*(3), 380–406.

Farnell, A. (2023). False Start? An Analysis of NFL Penalties With and Without Crowds. *Journal of Sports Economics*, *24*(6), 695–716.

Faruk, A. (2018). The Comparison of Proportional Hazards and Accelerated Failure Time Models in Analyzing the First Birth Interval Survival Data. *Journal of Physics: Conference Series*, *974*(1), 012008.

Ferraresi, M., & Gucciardi, G. (2021). Who Chokes on a Penalty Kick? Social Environment and Individual Performance During Covid-19 Times. *Economics Letters*, *203*, 109868.

Ferraresi, M., & Gucciardi, G. (2023). Team Performance and the Perception of Being Observed: Experimental Evidence from Top-Level Professional Football. *German Economic Review*, *24*(1), 1–31.

Figlio, D., Giuliano, P., Özek, U., & Sapienza, P. (2019). Long-Term Orientation and Educational Performance. *American Economic Journal: Economic Policy*, *11*(4), 272–309.

Fischer, K., & Haucap, J. (2022). Home Advantage in Professional Soccer and Betting Market Efficiency: The Role of Spectator Crowds. *Kyklos*, *75*(2), 294–316.

Flepp, R., & Franck, E. (2021). The Performance Effects of Wise and Unwise Managerial Dismissals. *Economic Inquiry*, *59*(1), 186–198.

Forbes. (2010). *The World's Most Valuable Teams And Athletes.* Retrieved from https://www.forbes.com/2010/07/20/most-valuable-athletes-and-teams-business-sports-sportsmoney-fifty-fifty.html

Forbes. (2011). *The World's 50 Most Valuable Sports Teams.* Retrieved from https://www.forbes.com/sites/kurtbadenhausen/2011/07/12/the-worlds-50-most-valuable-sports-teams/

Forbes. (2012). *Manchester United Tops The World's 50 Most Valuable Sports Teams.* Retrieved from https://www.forbes.com/sites/kurtbadenhausen/

2012/07/16/manchester-united-tops-the-worlds-50-most-valuable-sports-teams/

Foroughi, B., Gholipour, H. F., McDonald, H., & Jafarzadeh, B. (2018). Does National Culture Influence the Turnover Frequency of National Football Coaches? A Macro-Level Analysis. *International Journal of Sports Science & Coaching*, *13*(6), 902–911.

Fransen, K., Steffens, N. K., Haslam, S. A., Vanbeselaere, N., Vande Broek, G., & Boen, F. (2016). We Will Be Champions: Leaders' Confidence in 'Us' Inspires Team Members' Team Confidence and Performance. *Scandinavian Journal of Medicine & Science in Sports*, *26*(12), 1455–1469.

Fransen, K., Vanbeselaere, N., De Cuyper, B., Vande Broek, G., & Boen, F. (2014). The myth of the team captain as principal leader: extending the athlete leadership classification within sport teams. *Journal of Sports Sciences*, *32*(14), 1389–1397.

Fransen, K., Van Puyenbroeck, S., Loughead, T. M., Vanbeselaere, N., De Cuyper, B., Vande Broek, G., & Boen, F. (2015). Who Takes the Lead? Social Network Analysis as a Pioneering Tool to Investigate Shared Leadership Within Sports Teams. *Social Networks*, *43*, 28–38.

Frick, B., Barros, C. P., & Prinz, J. (2010). Analysing Head Coach Dismissals in the German "Bundesliga" With a Mixed Logit Approach. *European Journal of Operational Research*, *200*(1), 151–159.

Frick, B., & Simmons, R. (2008). The impact of managerial quality on organizational performance: evidence from German soccer. *Managerial and Decision Economics*, *29*(7), 593–600.

Fuchs, A., & Gehring, K. (2017). The Home Bias in Sovereign Ratings. *Journal of the European Economic Association*, *15*(6), 1386–1423.

Galor, O., & Özak, (2016). The Agricultural Origins of Time Preference. *American Economic Review*, *106*(10), 3064–3103.

Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism Under Social Pressure. *The Review of Economics and Statistics*, *87*(2), 208–216.

Ghimire, S., Ehrlich, J. A., & Sanders, S. D. (2020). Measuring Individual Worker Output in a Complementary Team Setting: Does Regularized Adjusted Plus Minus Isolate Individual NBA Player Contributions? *PLOS ONE*, *15*(8), e0237920.

Goldin, C., & Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *The American Economic Review*, *90*(4), 715–741.

Gould, E. D., & Winter, E. (2009). Interactions between Workers and the Technology of Production: Evidence from Professional Baseball. *The Review of Economics and Statistics*, *91*(1), 188–200.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232.

Grusky, O. (1963). Managerial Succession and Organizational Effectiveness. *American Journal of Sociology*, *69*(1), 21–31.

Guerrero-Calderón, B., Owen, A., Morcillo, J. A., & Castillo-Rodríguez, A. (2021). How Does the Mid-Season Coach Change Affect Physical Performance on Top Soccer Players? *Physiology & Behavior*, *232*, 113328.

Guryan, J., Kroft, K., & Notowidigdo, M. J. (2009). Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments. *American Economic Journal: Applied Economics*, *1*(4), 34–68.

Hart, O., & Holmström, B. (1987). The Theory of Contracts. In T. F. Bewley (Ed.), *Advances in Economic Theory: Fifth World Congress* (pp. 71–156). Cambridge University Press.

Herbst, D., & Mas, A. (2015). Peer Effects on Worker Output in the Laboratory Generalize to the Field. *Science*, *350*(6260), 545–549.

Hermalin, B. E. (1998). Toward an Economic Theory of Leadership: Leading by Example. *The American Economic Review*, *88*(5), 1188–1206.

Hermalin, B. E. (2013). 11. Leadership and Corporate Culture. In *The Handbook of Organizational Economics* (pp. 432–478). Princeton University Press.

Herrmann, M. A., & Rockoff, J. E. (2012). Worker Absence and Productivity: Evidence from Teaching. *Journal of Labor Economics*, *30*(4), 749–782.

Hill, N., & Remer, M. (2020). Race and Employment Outcomes: Evidence from NBA Coaches. *Economic Inquiry*, *58*(3), 1469–1486.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780.

Hodler, R., & Raschky, P. A. (2014). Regional Favoritism. *The Quarterly Journal of Economics*, *129*(2), 995–1033.

Hoey, S., Peeters, T., & van Ours, J. C. (2023). The Impact of Absent Co-workers on Productivity in Teams. *Labour Economics*, *83*, 102400.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. SAGE Publications.

Holmstrom, B., & Milgrom, P. (1991). Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *The Journal of Law, Economics, and Organization*, *7*, 24–52.

Hughes, M., Hughes, P., Mellahi, K., & Guermat, C. (2010). Short-term versus Long-term Impact of Managers: Evidence from the Football Industry. *British Journal of Management*, *21*(2), 571–589.

Humphreys, B. R., & Johnson, C. (2020). The Effect of Superstars on Game Attendance: Evidence From the NBA. *Journal of Sports Economics*, *21*(2), 152–175.

Hübner, M., & Vannoorenberghe, G. (2015). Patience and Long-run Growth. *Economics Letters*, *137*, 163–167.

Ichniowski, C., & Preston, A. (2014). *Do Star Performers Produce More Stars? Peer Effects and Learning in Elite Teams* (Working Paper Series No. 20478). National Bureau of Economic Research.

IFAB. (2022). *Laws of the Game 2022-23.* Retrieved from https://downloads.theifab.com/downloads/laws-of-the-game-2022-23?l=en

Ingersoll, K., Malesky, E., & Saiegh, S. M. (2017). Heterogeneity and Team Performance: Evaluating the Effect of Cultural Diversity in the World's Top Soccer League. *Journal of Sports Analytics*, *3*(2), 67–92.

Jane, W.-J. (2015). Peer Effects and Individual Performance: Evidence From Swimming Competitions. *Journal of Sports Economics*, *16*(5), 531–539.

Jane, W.-J. (2016). The Effect of Star Quality on Attendance Demand: The Case of the National Basketball Association. *Journal of Sports Economics*, *17*(4), 396–417.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, *3*(4), 305–360.

Jiang, L. (2020). Splash with a Teammate: Peer effects in High-stakes Tournaments. *Journal of Economic Behavior & Organization*, *171*, 165–188.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292.

Kaplan, S. M. (2022). Putting a Price on Popularity: Evidence from Superstars in the National Basketball Association. *Economic Inquiry*, *60*(3), 1357–1381.

Kattuman, P., Loch, C., & Kurchian, C. (2019). Management Succession and Success in a Professional Soccer Team. *PLOS ONE*, *14*(3), e0212634.

Kingma, D. P., & Ba, J. (2015). *Adam: A Method for Stochastic Optimization.* arXiv.

Krumer, A., & Lechner, M. (2018). Midweek Effect on Soccer Performance: Evidence from the German Bundesliga. *Economic Inquiry*, *56*(1), 193–207.

Krumer, A., Otto, F., & Pawlowski, T. (2022). Nationalistic Bias Among International Experts: Evidence from Professional Ski Jumping. *The Scandinavian Journal of Economics*, *124*(1), 278–300.

Kuziemko, I., & Werker, E. (2006). How Much Is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations. *Journal of Political Economy*, *114*(5), 905–930.

Lackner, M. (2023). Effort and Risk-taking in Tournaments with Superstars – Evidence for Teams. *Applied Economics*, *55*(57), 6776–6792.

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.

Le Pargneux, A., & Zeitoun, H. (2023). Patience and Subjective Well-being. *Applied Economics Letters*, *30*(14), 1923–1929.

Madum, A. (2016). Managerial Turnover and Subsequent Firm Performance: Evidence from Danish Soccer Teams. *International Journal of Sport Finance*, *11*(1), 46–63.

Malmendier, U., & Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics*, *124*(4), 1593–1638.

Mao, E. (2023). The Incentive Effects of Tournaments and Peer Effects in Team Production: Evidence from Esports. *Journal of Sports Economics*, *24*(2), 174–192.

Mas, A., & Moretti, E. (2009). Peers at Work. *American Economic Review*, *99*(1), 112–145.

Mills, B. M. (2014). Social Pressure at the Plate: Inequality Aversion, Status, and Mere Exposure. *Managerial and Decision Economics*, *35*(6), 387–403.

Molodchik, M., Paklina, S., & Parshakov, P. (2021). Peer Effects on Individual Performance in a Team Sport. *Journal of Sports Economics*, *22*(5), 571–586.

Morgulev, E., & Galily, Y. (2019). Analysis of Time-wasting in English Premier League Football Matches: Evidence for Unethical Behavior in Final Minutes of Close Contests. *Journal of Behavioral and Experimental Economics*, *81*, 1–8.

Mourao, P. R., & Araújo, P. (2024). "It's the Average Cost, Mister!" – Explaining the Recurrent Dismissal of European Soccer Coaches. *Applied Economics*, *56*(23), 2769–2789.

Narita, K., Tena, J. D., & Detotto, C. (2023). Causal Inference With Observational Data: A Tutorial on Propensity Score Analysis. *The Leadership Quarterly*, *34*(3), 101678.

Neugart, M., & Richiardi, M. G. (2013). Sequential Teamwork in Competitive Environments: Theory and Evidence from Swimming Data. *European Economic Review*, *63*, 186–205.

Nevill, A. M., Balmer, N., & Mark Williams, A. (2002). The Influence of Crowd Noise and Experience upon Refereeing Decisions in Football. *Psychology of Sport and Exercise*, *3*(4), 261–272.

Nevill, A. M., Newell, S. M., & Gale, S. (1996). Factors Associated With Home Advantage in English and Scottish Soccer Matches. *Journal of Sports Sciences*, *14*(2), 181–186.

Nieminen, M. (2022). Cross-country Variation in Patience, Persistent Current Account Imbalances and the External Wealth of Nations. *Journal of International Money and Finance*, *121*, 102517.

Pazzona, M. (2022). Peer Interactions and Performance in a High-skilled Labour Market. *The Scandinavian Journal of Economics*, *124*(4), 1087–1116.

Pearce, C. L., Sims, H. P., Cox, J. F., Ball, G., Schnell, E., Smith, K. A., & Trevino, L. (2003). Transactors, Transformers and Beyond: A Multi-method Development of a Theoretical Typology of Leadership. *Journal of Management Development*, *22*(4), 273–307.

Pettersson-Lidbom, P., & Priks, M. (2010). Behavior Under Social Pressure: Empty Italian Stadiums and Referee Bias. *Economics Letters*, *108*(2), 212–214.

Pollard, R. (1986). Home Advantage in Soccer: A Retrospective Analysis. *Journal of Sports Sciences*, *4*(3), 237–248.

Pollard, R., & Pollard, G. (2005). Home Advantage in Soccer: A Review of Its Existence and Causes. *International Journal of Soccer and Science*, *3*(1), 28–38.

Ponce-Bordón, J. C., López-Gajardo, M. A., Fernández-Navarro, J., López del Campo, R., Resta, R., & García-Calvo, T. (2023). The Effect of Coach Dismissal on Team Performance and Match Physical Demands in Spanish Professional Soccer Leagues. *International Journal of Performance Analysis in Sport*, 1–17.

Ponzo, M., & Scoppa, V. (2018). Does the Home Advantage Depend on Crowd Support? Evidence From Same-Stadium Derbies. *Journal of Sports Economics*, *19*(4), 562–582.

Pope, B. R., & Pope, N. G. (2015). Own-Nationality Bias: Evidence from UEFA Champions League Football Referees. *Economic Inquiry*, *53*(2), 1292–1304.

Pope, D. G., Price, J., & Wolfers, J. (2018). Awareness Reduces Racial Bias. *Management Science*, *64*(11), 4988–4995.

Price, J., & Wolfers, J. (2010). Racial Discrimination Among NBA Referees. *The Quarterly Journal of Economics*, *125*(4), 1859–1887.

Principe, F., & van Ours, J. C. (2022). Racial Bias in Newspaper Ratings of Professional Football Players. *European Economic Review*, *141*, 103980.

Rabin, M. (1998). Psychology and Economics. *Journal of Economic Literature*, *36*(1), 11–46.

Radzimiński, , Padrón-Cabo, A., Modric, T., Andrzejewski, M., Versic, S., Chmura, P., … Konefał, M. (2022). The Effect of Mid-season Coach Turnover on Running Match Performance and Match Outcome in Professional Soccer Players. *Scientific Reports*, *12*(1), 10680.

Reade, J. J., Schreyer, D., & Singleton, C. (2022). Eliminating Supportive Crowds Reduces Referee Bias. *Economic Inquiry*, *60*(3), 1416–1436.

Reilly, P., Solow, J. L., & von Allmen, P. (2023). When the Stars Are Out: The Impact of Missed Games on NBA Television Audiences. *Journal of Sports Economics*, *24*(7), 877–902.

Rickman, N., & Witt, R. (2008). Favouritism and Financial Incentives: A Natural Experiment. *Economica*, *75*(298), 296–309.

Rosen, S. (1981). The Economics of Superstars. *The American Economic Review*, *71*(5), 845–858.

Santos, F., Strachan, L., Gould, D., Pereira, P., & Machado, C. (2019). The Role of Team Captains in Integrating Positive Teammate Psychological Development in High-Performance Sport. *The Sport Psychologist*, *33*(1), 1–11.

Schreyer, D., & Singleton, C. (2023). *Cristiano of Arabia: Did Ronaldo Increase Saudi Pro League Attendances?* SSRN Electronic Journal.

Scoppa, V. (2008). Are Subjective Evaluations Biased by Social Factors or Connections? An Econometric Analysis of Soccer Referee Decisions. *Empirical Economics*, *35*(1), 123–140.

Scoppa, V. (2021). Social Pressure in the Stadiums: Do Agents Change Behavior Without Crowd Support? *Journal of Economic Psychology*, *82*, 102344.

Slade, P., & Tolhurst, T. (2019). Job Security and Risk-Taking: Theory and Evidence From Professional Football. *Southern Economic Journal*, *85*(3), 899–918.

Soebbing, B. P., Wicker, P., & Weimar, D. (2015). The Impact of Leadership Changes on Expectations of Organizational Performance. *Journal of Sport Management*, *29*(5), 485–497.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958.

Sunde, U., Dohmen, T., Enke, B., Falk, A., Huffman, D., & Meyerheim, G. (2022). Patience and Comparative Development. *The Review of Economic Studies*, *89*(5), 2806–2840.

Sutter, M., & Kocher, M. G. (2004). Favoritism of Agents – The Case of Referees' Home Bias. *Journal of Economic Psychology*, *25*(4), 461–469.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., & Trautmann, S. T. (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. *American Economic Review*, *103*(1), 510–531.

Szymanski, M., Fitzsimmons, S. R., & Danis, W. M. (2019). Multicultural Managers and Competitive Advantage: Evidence from Elite Football Teams. *International Business Review*, *28*(2), 305–315.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *Journal of Health Economics*, *27*(3), 531–543.

Thaler, R. H. (2015). *Misbehaving: The Making of Behavioral Economics*. W.W. Norton & Co.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061.

van Ours, J. C., & van Tuijl, M. A. (2016). In-Season Head-Coach Dismissals and the Performance of Professional Football Teams. *Economic Inquiry*, *54*(1), 591–604.

Waldinger, F. (2012). Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany. *The Review of Economic Studies*, *79*(2), 838–861.

Wang, M., Rieger, M. O., & Hens, T. (2016). How Time Preferences Differ: Evidence from 53 Countries. *Journal of Economic Psychology*, *52*, 115–135.

Weber, M. (1930). *The Protestant Ethic and the Spirit of Capitalism*. Unwin Hyman.

Wegelin, P., Orlowski, J., & Dietl, H. M. (2022). The Importance of High Performing Team Members in Complex Team Work: Results from Quasi-experiments in Professional Team Sports. *Economic Inquiry*, *60*(3), 1296–1310.

Wills, G., Tacon, R., & Addesa, F. (2022). Uncertainty of Outcome, Team Quality or Star Players? What Drives TV Audience Demand for UEFA Champions League Football? *European Sport Management Quarterly*, *22*(6), 876–894.

Yamane, S., & Hayashi, R. (2015). Peer Effects among Swimmers. *The Scandinavian Journal of Economics*, *117*(4), 1230–1255.

Zitzewitz, E. (2006). Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making. *Journal of Economics & Management Strategy*, *15*(1), 67–99.

Appendix A

Referee Bias in Football: Actual vs. Expected Additional Time

A.1 Robustness Checks

Table A.1 *Robustness Checks with Different Ball-in-Play Ratios*

BIP Ratio:	55%	60%	65%	70%
Score Difference	-23.91***	-25.05***	-26.19***	-27.32***
	(3.073)	(3.219)	(3.365)	(3.512)
Covid	17.29	17.42	17.56	17.71
	(15.73)	(16.48)	(17.23)	(17.97)
Attendance (1000)	-0.050	-0.056	-0.062	-0.068
	(0.271)	(0.284)	(0.297)	(0.309)
Weekday	2.568	2.601	2.434	2.67
	(3.596)	(3.768)	(3.939)	(4.110)
Round	-0.233	-0.234	-0.234	-0.235
	(0.192)	(0.193)	(0.194)	(0.195)
Quality Ratio	-0.26	-0.261	-0.262	-0.263
	(0.961)	(0.999)	(1.033)	(1.067)
Score Difference × Covid	19.48**	21.51**	23.53**	25.55**
	(6.735)	(7.056)	(7.377)	(7.697)
Fixed-effects				
Home Team ×*Season*	Yes	Yes	Yes	Yes
Away Team ×*Season*	Yes	Yes	Yes	Yes
Referee	Yes	Yes	Yes	Yes
League	Yes	Yes	Yes	Yes
Observations	3,491	3,491	3,491	3,491
R^2	0.516	0.516	0.516	0.516

Note: Data sources: Opta & Football-Reference. The dependent variable is the difference between actual and expected additional time(in seconds) at the end of the 90$^{\text{th}}$ minute of matches that ended with a 1-goal difference as in Garicano et al. (2005). Four-way clustered (Home Team-Season & Away Team-Season & Referee & League) robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

Table A.2 *Robustness Checks with Traditional Method*

Dependent Variable:	Actual Additional Time		
Model:	(1)	(2)	(3)
Score Difference	-17.31***	-16.28***	-22.98***
	(3.198)	(2.961)	(3.387)
Ball-in-Play	-0.111***	-0.110***	-0.111***
	(0.011)	(0.010)	(0.010)
Covid		28.51	18.53
		(20.70)	(21.08)
Attendance (1000)		-0.077	-0.033
		(0.379)	(0.364)
Weekday		0.886	0.837
		(3.769)	(3.779)
Round		-0.245	-0.227
		(0.233)	(0.230)
Quality Ratio		-1.294	-1.317
		(0.770)	(0.783)
Score Difference × Covid			20.59**
			(5.448)
Home Team ×*Season*	Yes	Yes	Yes
Away Team ×*Season*	Yes	Yes	Yes
Referee	Yes	Yes	Yes
League	Yes	Yes	Yes
Observations	3,550	3,491	3,491
R^2	0.507	0.520	0.522

Note: Data sources: Opta & Football-Reference. The dependent variable is the actual additional time(in seconds) played in the matches that 90^{th} minutes ended with a 1-goal difference as in Garicano et al. (2005). Ball-in-Play is used as a control for time lost. The difference in observation counts between models is due to missing data for attendance. Four-way clustered (Home Team-Season & Away Team-Season & Referee & Referee) robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

A.2 Magnitude of Bias by Leagues

Table A.3 *Examination of Bias by Leagues*

Dependent Variable:	Actual - Expected Additional Time					
Model:	(Turkish)	(English)	(French)	(German)	(Spanish)	(Italian)
Score Difference	-11.55	-20.82**	-23.64***	-13.59**	-26.37***	-16.59**
	(10.16)	(8.388)	(5.975)	(6.050)	(8.610)	(8.257)
Covid	28.16	-42.58		-44.15	85.96***	39.30*
	(23.67)	(30.74)		(30.95)	(17.04)	(19.95)
Attendance (1000)	0.3827	-1.775**	1.693*	-0.2166	-0.1731	0.1769
	(1.418)	(0.7053)	(0.9370)	(0.4497)	(0.3817)	(0.4179)
Weekday	-4.571	-1.031	16.91*	6.246	-7.825	0.3564
	(11.38)	(8.963)	(9.631)	(10.61)	(12.13)	(7.853)
Round	-0.9968**	0.0754	-0.2770	-0.1214	-0.1266	-0.4101
	(0.3806)	(0.3479)	(0.2541)	(0.4206)	(0.4062)	(0.3970)
Quality Ratio	3.124	-2.034	0.4721	-2.981	-0.4835	2.157
	(4.695)	(1.555)	(2.794)	(2.180)	(2.437)	(3.191)
Fixed-effects						
Home Team × Season	Yes	Yes	Yes	Yes	Yes	Yes
Away Team × Season	Yes	Yes	Yes	Yes	Yes	Yes
Referee	Yes	Yes	Yes	Yes	Yes	Yes
Observations	578	587	590	446	642	648
R^2	0.397	0.443	0.468	0.475	0.567	0.447

Note: Data sources: Opta & Football-Reference. The dependent variable is the difference between actual and expected additional time(in seconds) at the end of the 90^{th} minute of matches that ended with a 1-goal difference as in Garicano et al. (2005). The estimation for Covid-19 is not calculated for French League 1 as the league has been stopped during the pandemic. Three-way clustered (Home Team-Season & Away Team-Season & Referee) robust standard errors in parentheses. The variation in observation counts across leagues reflects the different number of matches played in each league that ended with a 1-goal difference. * p<0.1; ** p<0.05; *** p<0.01.

A.3 Experience and Quality of Referees

Table A.4 *Referee Experience, Quality and Bias*

Dependent Variable:	Actual - Expected Additional Time					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Score Difference	-20.38***	-20.55***	-20.30***	-53.17*	-20.04***	-22.37***
	(3.631)	(3.708)	(3.694)	(27.27)	(4.215)	(5.248)
Covid	26.94**	22.75*	22.84*	27.01**	22.78*	22.95*
	(12.62)	(12.62)	(12.60)	(12.59)	(12.61)	(12.62)
Attendance (1000)	-0.0800	-0.2152	-0.2047	-0.0778	-0.2155	-0.2030
	(0.2445)	(0.2446)	(0.2430)	(0.2452)	(0.2443)	(0.2438)
Weekday	2.584	4.789	4.688	2.515	4.795	4.638
	(3.758)	(3.615)	(3.600)	(3.765)	(3.613)	(3.601)
Round	-0.2651*	-0.2535	-0.2585	-0.2652*	-0.2559	-0.2538
	(0.1594)	(0.1590)	(0.1590)	(0.1594)	(0.1598)	(0.1596)
Quality Ratio	-1.129	-0.9208	-1.016	-1.188	-0.9239	-1.002
	(1.023)	(1.014)	(1.012)	(1.015)	(1.014)	(1.014)
Referee Age	0.6708**			0.2165		
	(0.3359)			(0.5162)		
Champions League Referee		9.268**			10.38*	
		(3.716)			(5.564)	
FIFA Referee			9.204***			6.966
			(3.309)			(5.004)
Score Difference × Referee Age				0.8327		
				(0.6819)		
Score Difference × UCL Referee					-2.114	
					(7.954)	
Score Difference × FIFA Referee						4.096
						(6.934)
Fixed-effects						
Home Team × Season	Yes	Yes	Yes	Yes	Yes	Yes
Away Team × Season	Yes	Yes	Yes	Yes	Yes	Yes
League	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,363	3,445	3,445	3,363	3,445	3,445
R^2	0.459	0.460	0.460	0.460	0.460	0.460

Note: Data sources: Opta & Football-Reference. The dependent variable is the difference between actual and expected additional time (in seconds) at the end of the 90th minute of matches that ended with a 1-goal difference as in Garicano et al. (2005). UCL and FIFA Referees are experienced in the UEFA Champions League and holders of FIFA badges. Four-way clustered (Home Team-Season & Away Team-Season & League & Referee) robust standard errors in parentheses. The difference in observation counts between models is due to missing data for referee age. * p<0.1; ** p<0.05; *** p<0.01.

A.4 Additional Time Intervals and Magnitude of Bias

Table A.5 *Additional Time Intervals and Magnitude of Bias*

Bias	-3 to -2	-2 to -1	-1 to 0	0 to +1	+1 to +2	+2 to +3	+3 to +4	+4 to +5	+5 to +6	+6 to +7
(Intercept)	-4.91^{***}	-2.18^{***}	-1.20^{***}	-0.91^{***}	-1.55^{***}	-2.02^{***}	-3.53^{***}	-3.07^{***}	-3.62^{***}	-5.39^{***}
	(0.41)	(0.15)	(0.11)	(0.10)	(0.13)	(0.17)	(0.29)	(0.36)	(0.47)	(0.85)
Score Difference	0.64^{**}	0.41^{***}	0.19^{**}	-0.05	-0.23^{**}	-0.36^{***}	-0.31	-0.03	-0.59	-0.31
	(0.28)	(0.11)	(0.08)	(0.08)	(0.09)	(0.13)	(0.21)	(0.28)	(0.40)	(0.60)
Covid	-0.55	-0.23	0.03	0.15	0.19	-0.20	-0.24	-0.86^{**}	-0.08	0.03
	(0.38)	(0.15)	(0.11)	(0.11)	(0.12)	(0.18)	(0.29)	(0.39)	(0.54)	(0.81)
Attendance (1000)	-0.01	-0.00	0.00	0.01^{**}	-0.00	-0.01	-0.01	-0.04^{***}	0.00	0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
Weekday	-0.19	-0.21	-0.27^{***}	0.10	0.24^{**}	0.07	0.08	0.33	0.40	-0.07
	(0.31)	(0.13)	(0.09)	(0.09)	(0.10)	(0.14)	(0.23)	(0.30)	(0.41)	(0.68)
Round	0.04^{***}	0.00	0.00	-0.01	0.00	-0.01	0.02	-0.03^{**}	-0.06^{***}	0.01
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)
Quality Ratio	0.02	0.01	-0.02	-0.00	-0.01	0.03	0.02	0.06	-0.08	-0.21
	(0.05)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.04)	(0.05)	(0.10)	(0.22)
AIC	611.82	2488.70	3934.62	4224.42	3281.61	1983.71	921.89	545.36	335.11	171.67
BIC	654.91	2531.78	3977.71	4267.50	3324.70	2026.80	964.98	588.44	378.20	214.76
Log Likelihood	-298.91	-1237.35	-1960.31	-2105.21	-1633.81	-984.86	-453.95	-265.68	-160.56	-78.84
Deviance	597.82	2474.70	3920.62	4210.42	3267.61	1969.71	907.89	531.36	321.11	157.67
Observations	3,481	3,481	3,481	3,481	3,481	3,481	3,481	3,481	3,481	3,481

Note: Data sources: Opta & Football-Reference. The dependent variables are dummy variables for 1-minute intervals of the differences between actual and expected additional time(in minutes) in the matches where any team leads by one goal difference as in Garicano et al. (2005). The estimates indicate that when the home team is ahead, there is an increased likelihood that the match will be concluded earlier, while if the home team is trailing, the match tends to be extended further. Robust standard errors are in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

Appendix B

Captains vs. All-Stars: Who Makes Better Leaders?

B.1 Robustness Checks for Team Quality Measures

Table B.1 *Robustness Checks with Betting Odds: Leader Type and Performance*

Dependent Variable:	Real Plus-Minus		
Model:	(1)	(2)	(3)
Both (Captain & All-Star)	2.937***	0.667***	0.492***
	(0.2603)	(0.2084)	(0.1663)
Captain Only	0.2980***	-0.414***	-0.379***
	(0.1811)	(0.1593)	(0.1142)
Allstar Only	1.705***	0.364	0.101
	(0.3205)	(0.2271)	(0.1995)
Playoff		0.265***	-0.945***
		(0.1025)	(0.0946)
Minutes		0.109***	0.121***
		(0.0045)	(0.0044)
Salary (10 Million USD)		0.204**	−0.0005
		(0.0929)	(0.0807)
Odds Ratio		1.28***	0.949***
		(0.0362)	(0.0341)
Age		0.431***	0.236**
		(0.1161)	(0.0924)
Age^2		-0.007***	-0.004**
		(0.0028)	(0.0025)
Fixed-effects			
Player Position	No	No	Yes
Team $\times$*Season*	No	No	Yes
Opponent Team $\times$*Season*	No	No	Yes
Observations	344,549	288,226	288,226
R^2	0.008	0.069	0.084

Note: Data Sources: ESPN & OddsPortal. Player-level clustered robust standard errors in parentheses. The variation in observation counts across models is due to missing data for betting odds before 2008. * p<0.1; ** p<0.05; *** p<0.01.

Table B.2 *Robustness Check with Betting Odds: Score Difference, Game Result and Injury of Key Players (Home Teams)*

Dependent Variables:	Score (1)	Score (2)	Score (3)	Result (1)	Result (2)	Result (3)
Injury of Both	-1.57^{***}	-1.46^{***}	-1.80^{***}	-0.19^{***}	-0.18^{***}	-0.25^{***}
	(0.20)	(0.19)	(0.30)	(0.03)	(0.03)	(0.06)
Injury of Only Captain	-1.59^{***}	-0.76^{***}	-0.13	-0.20^{***}	-0.06^{**}	0.02
	(0.17)	(0.16)	(0.28)	(0.03)	(0.03)	(0.05)
Injury of Only All-Star	0.15	-0.17	-0.50	0.04	-0.01	-0.07
	(0.18)	(0.17)	(0.29)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Both	1.20^{***}	0.94^{***}	1.61^{***}	0.13^{***}	0.10^{***}	0.21^{***}
	(0.20)	(0.19)	(0.25)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Only Captain	1.28^{***}	0.38^{**}	0.54^{***}	0.18^{***}	0.03	0.07
	(0.17)	(0.16)	(0.22)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Only All-Star	-0.64^{***}	-0.33	0.10	-0.07^{**}	-0.02	0.04
	(0.18)	(0.17)	(0.29)	(0.03)	(0.03)	(0.05)
Play-off		2.16^{***}	0.54		0.25^{***}	0.01
		(0.32)	(0.35)		(0.05)	(0.06)
Odds Ratio		2.91^{***}	-0.47^{***}		0.70^{***}	-0.06^{***}
		(0.05)	(0.09)		(0.02)	(0.02)
Fixed-effects						
Team × Season	No	No	Yes	No	No	Yes
Opponent Team × Season	No	No	Yes	No	No	Yes
Observations	25562	25549	25549	25562	25549	25549
R^2	0.01	0.11	0.25			
Deviance				34368.13	31523.32	28779.17
Log Likelihood				-17184.06	-15761.66	-14389.59
Pseudo R^2				0.01	0.09	0.10

Note: Data source: ESPN & OddsPortal. The first three columns show OLS estimations where the dependent variable is the score difference. The remaining three columns show logistic regression where the dependent variable is the match result. Odds Ratio is the ratio between the opponent's odds over team's odds. Therefore, the bigger Odds Ratio, the better the team than the opponent. Game-level clustered robust standard errors in parentheses. The decrease in observations from Models (1) and first models to the other models is due to missing data for the Elo Difference variable. * p<0.1; ** p<0.05; *** p<0.01.

B.2 Additional Estimations for Leaders and Others

Table B.3 *Minutes Played and Scoring Rates in Games*

Dependent Variables:	Minutes Played	Free Throw %	Field Goals %	3-Point Field Goal %
Model:	(1)	(2)	(3)	(4)
Both	6.484***	0.0043	-0.0144**	-0.0313***
	(0.7165)	(0.0104)	(0.0051)	(0.0079)
Captain Only	4.557***	-0.0042	-0.0054	-0.0111**
	(0.7986)	(0.0069)	(0.0044)	(0.0033)
All-Star Only	6.223***	-0.0018	-0.0067*	-0.0277***
	(0.4883)	(0.0159)	(0.0032)	(0.0041)
Playoff	0.0584	-0.0117***	-0.0121***	-0.0141***
	(0.2228)	(0.0030)	(0.0012)	(0.0020)
Real Plus Minus	0.0971***			
	(0.0033)			
Salary (10 Million USD)	5.6***	0.0089	−0.0073***	−0.0085**
	(0.4980)	(0.0070)	(0.0013)	(0.0025)
Elo Difference	-0.0026***	4.11×10^{-6}	-1.55×10^{-5}*	-5.28×10^{-5}***
	(0.0004)	(9.85×10^{-6})	(7.6×10^{-6})	(9.22×10^{-6})
Age	0.8447	0.0143**	0.0010	0.0082**
	(0.5660)	(0.0051)	(0.0024)	(0.0032)
Age^2	-0.0200*	-0.0002**	-2.95×10^{-5}	-9.52×10^{-5}
	(0.0091)	(8.37×10^{-5})	(4.12×10^{-5})	(5.7×10^{-5})
Minutes Played		0.0026***	0.0039***	0.0038***
		(0.0002)	(0.0002)	(0.0002)
Fixed-effects				
Player Position	Yes	Yes	Yes	Yes
Team × Season	Yes	Yes	Yes	Yes
Opponent Team × Season	Yes	Yes	Yes	Yes
Observations	288,226	213,278	346,074	230,452
R^2	0.284	0.0517	0.0608	0.0221

Note: Data Sources: ESPN & FiveThirtyEight. Player-level clustered robust standard errors in parentheses. The variation in observation counts across models is due to differences in observation number for each dependent variable. * p<0.1; ** p<0.05; *** p<0.01.

Table B.4 *Minutes Played and Scoring Rates in Games*

Dependent Variables:	Defensive Rebounds	Steals	Fouls Committed	Turnovers	Assists	Fouls Suffered
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Both	0.5855***	0.1253**	-0.1497*	0.5240***	0.8043***	1.934***
	(0.1437)	(0.0391)	(0.0685)	(0.0520)	(0.1798)	(0.1673)
Captain Only	0.1152	0.0430	0.0177	0.1012**	0.0891	0.1976**
	(0.1241)	(0.0315)	(0.0295)	(0.0327)	(0.1199)	(0.0599)
Allstar Only	0.3031**	0.0185	-0.1135*	0.3640***	0.5828**	0.7657***
	(0.1131)	(0.0225)	(0.0505)	(0.0445)	(0.2023)	(0.1160)
Playoff	-0.0250	-0.0495***	0.1777***	-0.0765***	-0.2124***	0.1508***
	(0.0428)	(0.0100)	(0.0112)	(0.0123)	(0.0428)	(0.0181)
Minutes Played	0.1315***	0.0317***	0.0578***	0.0465***	0.0912***	0.0996***
	(0.0165)	(0.0013)	(0.0013)	(0.0030)	(0.0162)	(0.0033)
Salary (10 Million USD)	0.2600***	-0.0167	-0.0767**	0.2460***	0.4173***	0.4211***
	(0.0420)	(0.0185)	(0.0323)	(0.0312)	(0.0885)	(0.0726)
Elo Difference	-0.0002*	5.47×10^{-6}	-0.0002**	-3.05×10^{-5}*	-0.0003***	0.0002***
	(0.0001)	(3.11×10^{-5})	(6.52×10^{-5})	(1.63×10^{-5})	(8.05×10^{-5})	(5.09×10^{-5})
Age	-0.1865***	0.0136	-0.0214	-0.0704*	0.0425	-0.0156
	(0.0393)	(0.0097)	(0.0157)	(0.0355)	(0.0787)	(0.0356)
Age^2	0.0028***	-0.0003*	0.0003	0.0007	-0.0009	-0.0014*
	(0.0007)	(0.0002)	(0.0003)	(0.0006)	(0.0015)	(0.0006)
Fixed-effects						
Player Position	Yes	Yes	Yes	Yes	Yes	Yes
Team × Season	Yes	Yes	Yes	Yes	Yes	Yes
Opponent Team ×Season	Yes	Yes	Yes	Yes	Yes	Yes
Observations	363,143	363,143	363,143	363,143	363,143	363,143
R^2	0.447	0.166	0.212	0.277	0.448	0.327

Note: Data Sources: ESPN & FiveThirtyEight. Player level clustered robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

B.3 Adjusted Plus-Minus (APM) Methodology

Adjusted Plus-Minus (APM) is a statistical model used to estimate the individual contribution of a player to their team's point differential, adjusting for the quality of teammates and opponents. In play-by-play analysis, APM is applied to each play to capture a player's impact at every play of the game.

For each play, the point differential (PD) can be modeled as[1]:

$$PD_p = \sum_{i=1}^{N} \beta_i X_{i,p} + \epsilon_p, \tag{B.1}$$

where:

- PD_p is the change in the point differential that occurs as a result of play p. Positive values indicate a net score increase for the team and negative values indicate a net score increase for the opponent.
- $X_{i,p}$ is a dummy variable for player i during play p ($X_{i,p} = 1$ if player i is on the court during play p, and $X_{i,p} = 0$ otherwise).
- β_i is the estimated contribution of player i to the point differential.
- ϵ_p is the error term for play p which captures unobserved factors affecting the point differential during that play.

This model is run over each play and the coefficients β_i represent the player's adjusted impact on team production over time.

B.4 LSTM Memory and Predicted Real Plus-Minus

B.4.1 LSTM Memory Specification

The LSTM model architecture consisted of an input layer, two LSTM layers with 128 units each, a dropout layer (rate = 0.2) for regularisation, and a dense output layer. In

[1] Further details on the Adjusted Plus-Minus by the creator can be found on the website at https://www.82games.com/ilardi1.htm.

neural networks, a layer is a collection of nodes that process input data and pass the results to the next layer. The LSTM layers are specialised for processing sequential data, while the dense output layer produces the final prediction. The dropout layer helps to 'forget' or discount historical performance, allowing the model to focus more on recent plays rather than being overly influenced by past performance. The rate of 0.2 was chosen to provide a balance between retaining enough information for learning and introducing sufficient variability to capture the changes (Srivastava et al., 2014) in player performance within a game, where a player's effectiveness can vary due to factors such as fatigue and number of personal fouls committed.

Mathematically, the model can be represented as:

$$h_t = LSTM(x_t, h_{t-1}) \tag{B.2}$$

$$RPM_t = Wh_t + b \tag{B.3}$$

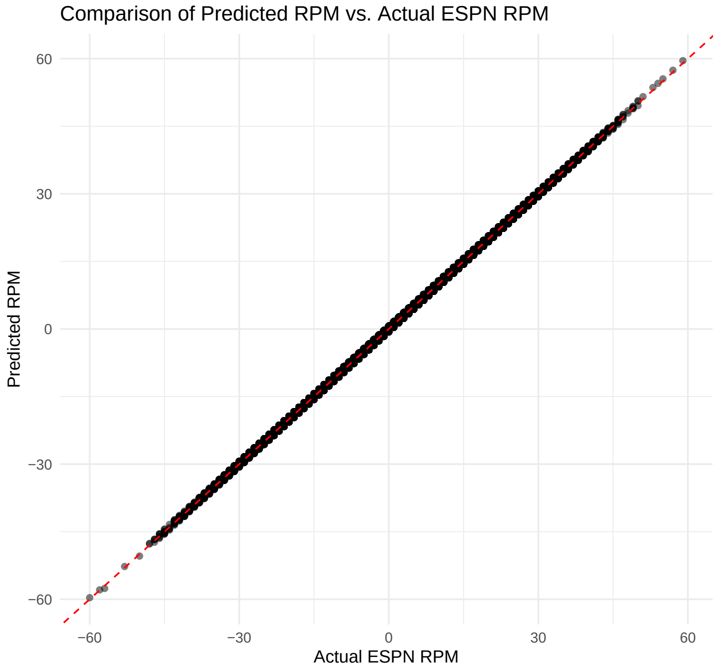
where x_t represents the input features at time step t, including player ID, current game statistics (points, rebounds, assists, etc.), team score differential, time remaining, lineup composition, and historical player performance statistics. The hidden state h_t carries information from previous time steps (past plays) which allows the model to consider both recent and long-term patterns in player performance. W and b are learnable parameters that the model adjusts during training to optimise its predictions.

The Adam optimiser was employed with a learning rate of 0.001 which is a standard value that balances the convergence speed and stability of the training process. Adam is a widely used optimisation algorithm that adapts the learning rates of parameters based on estimates of the first and second moments of the gradients (i.e., the mean and uncentred variance), which helps in efficient and effective training (Kingma & Ba, 2015). The Mean Squared Error (MSE) was used as the loss function, which measures the average of the squared differences between predicted and actual values. This helps the model to minimise prediction errors during training.

Model performance was validated by comparing end-of-game RPM predictions to the actual RPM values provided by ESPN. Figure B.1 illustrates a scatter plot of predicted versus actual RPM values and demonstrates a strong correlation ($R^2 = 0.98$).

B.4.2 Validation of LSTM-based Prediction with ESPN Real Plus-Minus

Fig. B.1 *Predicted and ESPN Real Plus-Minus*



B.5 Distribution of Absence of Leaders

Table B.5 *Observations Pre and Post Treatment by Absence Reason, Leader Type and Season*

Reason	6^{th} Foul						Injuries					
	Allstar		Captain		Both		Allstar		Captain		Both	
Season	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
2002	210175	1709	210466	1418	210791	1093	203572	8312	202097	9787	205814	6070
2003	434595	2921	433050	4466	435013	2503	418635	18881	421449	16067	423427	14089
2004	531083	4134	531500	3717	533008	2209	520310	14907	519744	15473	529615	5602
2005	555162	3765	552709	6218	556148	2779	536095	22832	537716	21211	548415	10512
2006	556928	3202	554070	6060	557630	2500	540975	19155	538326	21804	547619	12511
2007	558077	3291	556006	5362	558769	2599	539805	21563	544828	16540	549575	11793
2008	560568	3064	559517	4115	561160	2472	546897	16735	551892	11740	553990	9642
2009	558323	3806	556887	5242	559130	2999	554848	7281	541802	20327	555109	7020
2010	563065	3242	562124	4183	563799	2508	555061	11246	555250	11057	561736	4571
2011	564764	1819	562886	3697	565106	1477	550489	16094	552661	13922	557571	9012
2012	461206	1071	460716	1561	461323	954	436643	25634	449328	12949	454896	7381
2013	562605	1999	561796	2808	563029	1575	552047	12557	550143	14461	557897	6707
2014	578148	2621	576458	4311	578684	2085	557870	22899	561432	19337	568176	12593
2015	579305	1522	578427	2400	579368	1459	549608	31219	566570	14257	567526	13301
2016	584630	1583	584327	1886	584956	1257	574995	11218	580966	5247	577887	8326
2017	585294	1588	585224	1658	585716	1166	575959	10923	574033	12849	586209	673
2018	589526	1008	588825	1709	589551	983	577253	13281	580178	10356	585444	5090
2019	610624	2331	609720	3235	611271	1684	576141	36814	584246	28709	595727	17228
2020	537614	1305	536552	2367	538061	858	524419	14500	519135	19784	530813	8106
2021	529794	1276	529725	1345	530566	504	512083	18987	509196	21874	524790	6280

Note: Data source: Own calculations based on data from ESPN. 'Pre' and 'Post' columns show the number of observations before and after the treatment (player absence) respectively. 'Both' refers to players who are both All-Stars and team captains.

B.6 Distance of Two-Point Field Goals

Table B.6 *Distance of Two-Point Field Goal Attempts*

Dependent Variable:	Distance of Two Point Field Goal Attempts		
Model:	(1)	(2)	(3)
Both: Treatment × Post	-0.0370		
	(0.2988)		
Captain: Treatment × Post		-0.1456	
		(0.2807)	
All-Star: Treatment × Post			-0.7886
			(0.5478)
Home	-0.2784***	-0.2779***	-0.2781***
	(0.0144)	(0.0144)	(0.0144)
Score Difference	0.0024**	0.0024***	0.0024***
	(0.0009)	(0.0009)	(0.0009)
Period	-0.1521***	-0.1518***	-0.1535***
	(0.0104)	(0.0104)	(0.0104)
Fixed-effects			
Game	Yes	Yes	Yes
Player	Yes	Yes	Yes
Team	Yes	Yes	Yes
Opponent Team	Yes	Yes	Yes
Observations	1,760,377	1,760,377	1,760,377
R^2	0.141	0.141	0.141

Note: Data source: ESPN. Score Difference is $Team - Opponent$. Player-level clustered robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01.

B.7 Estimations for Away Teams

Table B.7 *Score Difference, Game Result and Injury of Key Players (Away Teams)*

Dependent Variable:	Score (1)	Score (2)	Score (3)	Result (1)	Result (2)	Result (3)
Injury of Both	-1.20^{***}	-0.87^{***}	-1.65^{***}	-0.13^{***}	-0.10^{**}	-0.22^{***}
	(0.20)	(0.19)	(0.25)	(0.03)	(0.03)	(0.05)
Injury of Only Captain	-1.28^{***}	-0.09	-0.57^{*}	-0.18^{***}	-0.01	-0.08
	(0.17)	(0.15)	(0.24)	(0.03)	(0.03)	(0.05)
Injury of Only All-Star	0.64^{***}	0.17	-0.12	0.07^{*}	0.00	-0.05
	(0.19)	(0.17)	(0.28)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Both	1.57^{***}	1.23^{***}	1.89^{***}	0.19^{***}	0.16^{***}	0.27^{***}
	(0.20)	(0.18)	(0.25)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Only Captain	1.59^{***}	0.30	0.17	0.20^{***}	0.02	-0.01
	(0.17)	(0.16)	(0.25)	(0.03)	(0.03)	(0.05)
Opponent's Injury of Only All-Star	-0.15	0.27	0.54^{*}	-0.04	0.01	0.08
	(0.19)	(0.17)	(0.28)	(0.03)	(0.03)	(0.05)
Play-off		-1.22^{***}	-0.74^{*}		-0.11^{*}	-0.03
		(0.32)	(0.35)		(0.05)	(0.06)
Elo Difference		0.04^{***}	-0.01^{***}		0.01^{***}	-0.00^{***}
		(0.00)	(0.00)		(0.00)	(0.00)
Fixed-effects						
Team × *Season*	No	No	Yes	No	No	Yes
Opponent × *Season*	No	No	Yes	No	No	Yes
Observations	25,562	25,549	25,549	25,562	25,549	25,549
R^2	0.01	0.16	0.25			
Deviance				34368.13	31001.57	28727.36
Log Likelihood				-17184.06	-15500.79	-14363.68
Pseudo R^2				0.01	0.10	0.10

Note: Data sources: ESPN & FiveThirtyEight. Game level clustered robust standard errors in parentheses. The decrease in observations from the first models to the other models is due to missing data for the Elo Difference variable. * p<0.1; ** p<0.05; *** p<0.01.

Appendix C

The Effect of Patience on Tenure of Managers: Evidence From Football

C.1 Descriptive Statistics

Table C.1 *Descriptive Statistics of Match-Level Data*

Statistic	N	Mean	St. Dev.	Min	Max
Manager-Match Related Statistics					
Match Week	363,473	16.572	10.481	1	70
Goals	358,373	1.347	1.238	0	16
Attendance	203,752	14,165.960	15,056.850	1	95,795
Points	363,473	1.459	1.294	0	3
Goal Difference	358,268	0.127	1.780	−16	16
Home	363,473	0.500	0.500	0	1
League Tier	363,473	1.125	0.357	1	4
Rank	363,473	5.587	4.457	1	38
Expected Points	363,473	1.373	0.452	0.075	2.883
Surprise	363,473	0.009	1.250	−2.741	2.823
Cumulative Surprise	363,473	6.083	24.601	−84.658	514.237
Team Promoted[a]	363,472	0.0003	0.016	0	1
Won Domestic Super Cup[a]	363,473	0.004	0.060	0	1
Won Domestic Cup[a]	363,473	0.004	0.059	0	1
Won Continental Cup 1[a]	363,473	0.0004	0.020	0	1
Won Continental Super Cup[a]	363,473	0.0005	0.022	0	1
Won Continental Cup 2[a]	363,473	0.0002	0.014	0	1
Won FIFA Club World Cup[a]	363,473	0.0005	0.022	0	1

Continued on next page

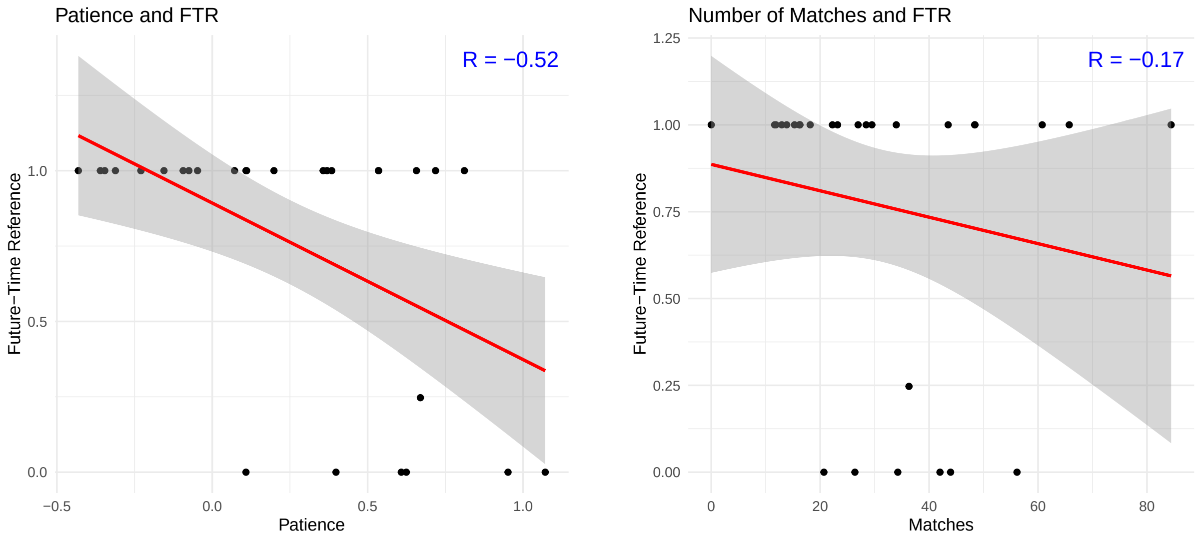
Table C.1 continued from previous page

Statistic	N	Mean	St. Dev.	Min	Max
Won Domestic League[a]	363,473	0.128	0.335	0	1
Dismissed	363,473	0.122	0.327	0	1
Dismissed during the Season	363,473	0.058	0.234	0	1
Age at Appointment	362,682	47.484	7.559	23	79
Foreign Manager	363,473	0.276	0.447	0	1
Country-Related Statistics					
Patience (GPS)	338,401	0.272	0.391	−0.381	1.071
Patience (Wang)	250,139	0.636	0.154	0.370	0.890
Long Term Orientation	363,473	49.606	17.576	5	100
Future-Term Reference (FTR)	307,974	0.708	0.445	0	1
Share of Protestants	338,401	0.153	0.199	0.000	0.629
Life Expectancy	363,473	79.686	3.600	62.649	83.794
Satisfaction with Life (Rank)	352,996	56.522	37.116	1	174
Mean Years of Schooling	338,401	10.605	1.845	5.387	13.424
Cognitive Skills	321,798	4.739	0.466	3.089	5.338
Democracy Index	338,401	9.054	2.027	0	10
Political Stability	363,473	0.332	0.625	−1.584	1.115
Subjective Institutions	229,071	48.418	11.322	26.386	76.074
Freedom Index	328,298	0.778	0.131	0.373	0.945
Social Support	328,298	0.896	0.067	0.511	0.948
Colonized	338,401	0.278	0.448	0	1
GDP per Capita[b]	356,989	36,949	20,539	2,238	93,446
Household Savings[b]	264,824	5.844	6.334	−8.255	40.702
Total Factor Productivity	338,401	0.783	0.156	0.352	1.041
Interest Rate	156,421	5.431	4.722	0.030	16.352
Temperature	332,980	11.791	5.897	−7.929	25.418
Precipitation	332,980	78.383	36.219	8.415	215.848
Unemployment	328,298	9.775	6.066	2.050	24.790
Pollution	328,298	17.762	12.939	6.628	88.169
FIFA Rank of Country	363,473	18.974	19.573	1	106
Corruption (Rank)	363,473	41.321	38.880	1	177
Happiness (Rank)	363,473	35.211	28.087	2	136
Rule of Law	355,217	0.692	0.132	0.299	0.889

Note: Data sources: Falk et al. (2018), Wang et al. (2016), Hofstede (2001), TransferMarkt & World Bank. [a] indicates "Last season with the manager", [b] indicates nominal USD.

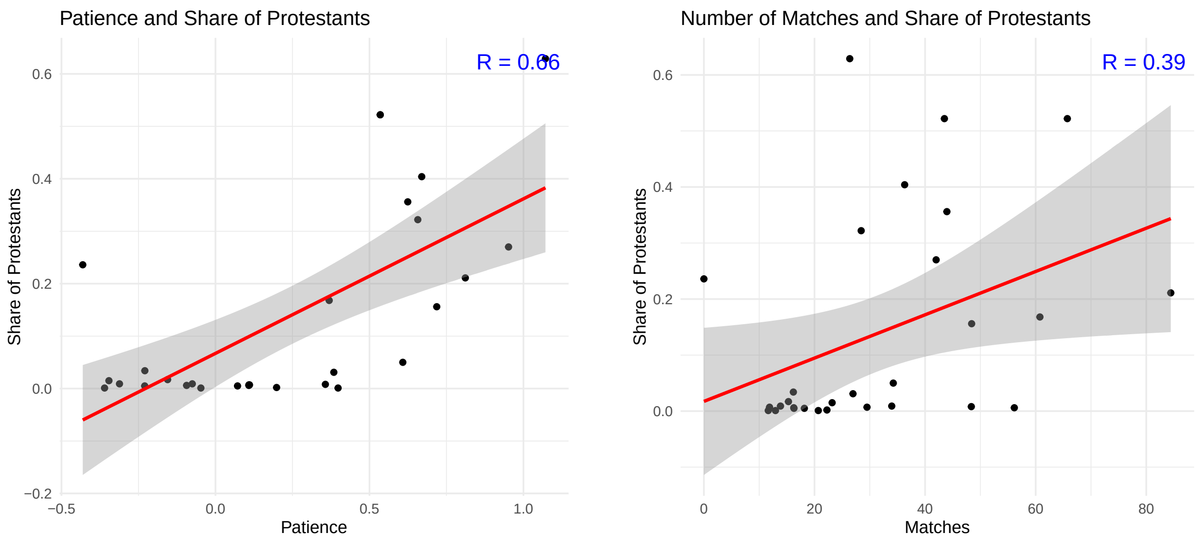
C.2 Validation of Instrumental Variables

Fig. C.1 *Scatter Plots of Patience, Number of Matches and Future-Time Reference*



Note: Data sources: Falk et al. (2018) & TransferMarkt. Red lines are LM curves. Shaded areas are 95% confidence intervals.

Fig. C.2 *Scatter Plots of Patience, Number of Matches and Share of Protestants*



Note: Data sources: Falk et al. (2018) & Le Pargneux & Zeitoun (2023). Red lines are LM curves. Shaded areas are 95% confidence intervals.

C.3 IV Results of Match-Level Data

Table C.2 *Patience and Tenure: Instrumental Variable Logistic Regression Results*

	Dismissal	
	2SPS	2SRI
Patience (Falk et al., 2018)	-20.214^{***}	-18.023^{***}
	(6.576)	(5.461)
Foreign	-0.493	-0.424
	(0.325)	(0.295)
Home	-0.018	-0.018
	(0.013)	(0.013)
League Tier	-0.824^{***}	-0.792^{***}
	(0.168)	(0.169)
Age at Appointment	0.039^{***}	0.038^{***}
	(0.013)	(0.013)
FIFA Rank	0.005	0.008
	(0.013)	(0.015)
Cumulative Surprise	-0.011^{*}	-0.011^{*}
	(0.006)	(0.006)
Promoted Last Year	3.253^{***}	3.236^{***}
	(0.553)	(0.545)
Won Domestic Super Cup	-0.516	-0.630
	(0.602)	(0.584)
Won Domestic Cup	-0.323	-0.485
	(0.662)	(0.679)
Won UEFA Champions League	0.731	0.422
	(0.564)	(0.637)
Won UEFA Super Cup	1.557^{**}	1.570^{**}
	(0.631)	(0.716)
Won Europa League	-2.029^{***}	-2.037^{***}
	(0.129)	(0.121)
League Winner Coach	0.577^{**}	0.585^{**}
	(0.278)	(0.280)
Patience (Residual)		-4.657^{***}
		(1.735)
Country-Specific Controls	Yes	Yes
Fixed-effects		
Manager	Yes	Yes
Observations	255234	255234
Deviance	102600.345	102334.020
Log Likelihood	-51300.172	-51167.010
Pseudo R^2	0.366	0.367

Data sources: Falk et al. (2018), TransferMarkt, World Bank. Match-Level data. Dependent variable is a binary variable being dismissed after the match. The instrumental variable is FTR. 2SPS: Two-Stage Predictor Substitution; 2SRI: Two-Stage Residual Inclusion. The cup and league variables are binary variables showing if the team won it with the same manager in the previous season. Country-Related controls include GDP per capita, inflation, unemployment, democracy index, and life expectancy among others. A full list of controls can be found in Table C.1 in the Appendix. Country-level clustered robust standard errors are in parentheses.
* p<0.1; ** p<0.05; *** p<0.01.

C.4 Results of Alternative Measures of Patience

Table C.3 *Patience and Tenure: IV Results of Alternative Patience Measures*

	Dismissal			
	2SPS	2SRI	2SPS	2SRI
Long Term Orientation (Hofstede, 2001)	-0.096^{***}	-0.065^{***}		
	(0.027)	(0.018)		
Patience (Wang et al., 2016)			-35.347^{***}	-33.074^{***}
			(10.783)	(9.908)
Foreign	-0.394	-0.324^{*}	-1.389^{***}	-1.367^{***}
	(0.255)	(0.190)	(0.517)	(0.473)
Home	-0.018	-0.018	-0.017	-0.017
	(0.013)	(0.013)	(0.014)	(0.014)
League Tier	-0.741^{***}	-0.726^{***}	-0.690^{***}	-0.695^{***}
	(0.160)	(0.165)	(0.187)	(0.188)
Age at Appointment	0.037^{***}	0.034^{**}	0.049^{***}	0.048^{***}
	(0.013)	(0.014)	(0.015)	(0.016)
FIFA Rank	0.018	0.009	0.026^{*}	0.028
	(0.014)	(0.010)	(0.015)	(0.018)
Cumulative Surprise	-0.011^{**}	-0.012^{**}	-0.009^{*}	-0.009^{*}
	(0.006)	(0.006)	(0.005)	(0.005)
Promoted Last Year	3.143^{***}	3.152^{***}	3.027^{***}	3.039^{***}
	(0.533)	(0.530)	(0.681)	(0.675)
Won Domestic Super Cup	-0.679	-0.550	-1.051	-1.050
	(0.601)	(0.684)	(0.917)	(0.925)
Won Domestic Cup	-0.551	-0.288	-1.074^{*}	-1.184^{*}
	(0.650)	(0.635)	(0.617)	(0.632)
Won Continental Cup 1	-0.033	0.146	-1.288^{*}	-1.417^{**}
	(0.585)	(0.674)	(0.666)	(0.671)
Won Continental Super Cup	1.208^{*}	0.847	1.291^{**}	1.386^{**}
	(0.638)	(0.722)	(0.590)	(0.687)
Won Continental Cup 2	-2.130^{***}	-2.013^{***}	-2.131^{***}	-2.136^{***}
	(0.136)	(0.057)	(0.078)	(0.094)
Won Domestic League	0.625^{**}	0.618^{**}	0.561^{*}	0.564^{*}
	(0.275)	(0.281)	(0.311)	(0.312)
League Winner Coach	0.045^{**}	0.089^{***}		
	(0.019)	(0.025)		
Long Term Orientation (Residual)		-0.100^{***}		
		(0.017)		
Patience (Residual)				5.728^{*}
				(3.162)
Country-Specific Controls	Yes	Yes	Yes	Yes
Fixed-effects				
Manager	Yes	Yes	Yes	Yes
Num. obs.	255234	255234	206707	206707
Deviance	102694.117	101948.543	85670.787	85481.843
Log Likelihood	-51347.058	-50974.271	-42835.394	-42740.921
Pseudo R^2	0.365	0.370	0.378	0.379

Data sources: Hofstede (2001), Wang et al. (2016), TransferMarkt, World Bank. Match-Level data. Dependent variable is a binary variable being dismissed after the match. The instrumental variable is FTR. 2SPS: Two-Stage Predictor Substitution; 2SRI: Two-Stage Residual Inclusion. The cup and league variables are binary variables showing if the team won it with the same manager in the previous season. Country-Related controls include GDP per capita, inflation, unemployment, democracy index, and life expectancy among others. A full list of controls can be found in Table C.1 in the Appendix. Country-level clustered robust standard errors are in parentheses. The variation in observation counts across models is due to differences in data availability for the alternative measures of patience. * p<0.1; ** p<0.05; *** p<0.01.

Table C.4 *Patience and Tenure: AFT Results of Alternative Measures of Patience*

	Time to Dismissal			
	(All)	(In-Season)	(All)	(In-Season)
Long Term Orientation (Hofstede, 2001)	0.094***	0.090***		
	(0.028)	(0.026)		
Patience (Wang et al., 2016)			1.684***	1.601***
			(0.098)	(0.144)
Foreign Manager	−0.165	−0.231	−0.368	−0.326
	(0.181)	(0.167)	(0.175)	(0.210)
Home	0.015	0.023	0.012	0.015
	(0.015)	(0.021)	(0.016)	(0.023)
League Tier	−0.293***	−0.298***	−0.144***	−0.162***
	(0.020)	(0.029)	(0.022)	(0.031)
Age at appointment	0.004	0.009	0.002	0.007
	(0.006)	(0.006)	(0.002)	(0.004)
FIFA Rank	−0.012	−0.014	−0.021	−0.020
	(0.017)	(0.019)	(0.018)	(0.019)
Cumulative Surprise	0.029***	0.036***	0.028***	0.033***
	(0.006)	(0.008)	(0.006)	(0.008)
Team Promoted	−1.823***	1.583*	−1.543***	1.577
	(0.227)	(0.892)	(0.548)	(1.247)
Won Domestic Super Cup	2.218***	2.126***	3.103***	2.594***
	(0.358)	(0.514)	(0.468)	(0.558)
Won Domestic Cup	2.071***	3.121***	2.774***	3.748***
	(0.345)	(0.755)	(0.424)	(0.896)
Won Continental Cup 1	2.736**	15.502	3.136**	15.569**
	(1.268)	(271.136)	(1.256)	(0.918)
Won Continental Super Cup	0.876	0.475	0.734*	0.650
	(0.761)	(1.305)	(0.315)	(1.267)
Won Continental Cup 2	4.488***	21.787***	4.518***	21.546***
	(1.236)	(0.913)	(1.238)	(0.846)
Won Domestic League	−0.301***	−0.522***	−0.212	−0.405*
	(0.019)	(0.022)	(0.132)	(0.182)
Country-Specific Controls	Yes	Yes	Yes	Yes
Observations	270,912	270,912	218,276	218,276
Log Likelihood	−197,544.000	−104,546.300	−174,925.900	−92,647.680
χ^2 (df = 22)	22,378.310***	11,295.350***	10,850.860***	5,910.167***

Data sources: Hofstede (2001), Wang et al. (2016), TransferMarkt, World Bank. Match-Level data. Dependent variable is the number of matches left to the dismissal. The cup and league variables are binary variables showing if the team won it with the same manager in the previous season. Country-Related controls include GDP per capita, inflation, unemployment, democracy index, and life expectancy among others. A full list of controls can be found in Table C.1 in the Appendix. Robust standard errors are in parentheses. The variation in observation counts across models is due to differences in data availability for the alternative measures of patience. * p<0.1; ** p<0.05; *** p<0.01.