Erasmus University Rotterdam
Erasmus School of Economics
Master Thesis Econometrics and Management Science

Evaluating the effect of cup participation on league performance: Evidence from European domestic cups

Jelle Willekes (609948)

Supervisor:	prof.dr. HD (Dinand) Webbink
Second assessor:	dr. W (Wendun) Wang
Date final version:	18th November 2024

Abstract

This paper investigates the impact of participating or winning in a domestic cup fixture on performance in the next league fixture in European football, focusing on four major domestic cup competitions: the FA Cup (England), DFB Pokal (Germany), KNVB Beker (Netherlands) and Taça de Portugal (Portugal). For estimating the causal effect of both winning and participating in a cup round on subsequent league performance, we use the randomness of the draws of cup fixtures an exogenous source of variation using a Two-Stage Least Squares (2SLS) instrumental variable approach. Our analysis shows that winning a cup fixture has a significant positive effect on league performance, particularly for teams from lower leagues and teams with a lower total market value. These teams appear to experience a boost in momentum after cup success, which translates into better performance in the next league fixture. However, for teams with higher total market value, the effect of winning on league performance is not statistically significant, suggesting that larger teams are more capable of managing fixture congestion. Moreover, we find no significant effect for participating in a cup fixture, regardless of the result, on performance in the next league fixture. This suggests that fixture congestion alone does not affect league performance. The findings are in line with prior research in international club competitions. This research adds to the existing literature on fixture congestion and football performance by showing how winning cup fixtures affects teams differently, depending on their market value and team size. The results provide useful insights for football clubs and managers, highlighting how smaller teams can strategically use cup momentum to their advantage in congested periods of a football season.

Contents

Chapter 1

Introduction

In modern football, fixture congestion is a frequently discussed topic, with both players and managers commenting on the negative effects on performance and player well-being. Top managers like José Mourinho, Jürgen Klopp, and Pep Guardiola have openly criticized increased fixture scheduling, arguing that crowded schedules give their teams a disadvantage, especially during critical phases of the season. Mourinho, during his time at Manchester United, emphasized how the congested calendar prevented his team from performing at their best (Fox Sports, 2022), while Klopp blamed both the Premier League and the TV networks for worsening the already crowded Liverpool schedule (The Guardian, 2022). Similarly, Guardiola expressed concerns about the impact on player health, warning that the demanding schedule is "too much" and puts player well-being at risk (Olley, 2020).

Players have also complained recently. Manchester City midfielder Rodri warned that players are close to a strike due to the congested calendar, citing its impact on both physical and mental well-being (Colman, 2023). Liverpool goalkeeper Alisson Becker criticized the new Champions League format for increasing the weight on players without considering their welfare. Alisson indicated that despite repeated calls from players, their opinions are often ignored, leading to extreme demands on their fitness and performance (Pearce, 2024).

In the context of fixture congestion, sports-specific research has explored the effects of travel and rest days on team performance. Oberhofer, Philippovich and Winner (2010) and Nichols (2014) analyzed these impacts, finding that travel fatigue had a minor effect, while rest days did not significantly influence performance, particularly in the National Basketball Association (NBA) and international men's football, as shown by Scoppa (2015). Additionally, Krumer and Lechner (2018) examined midweek league matches and observed that home advantage often decreased during these fixtures due to smaller crowds and lower perceived importance. Both home and away teams were also found to be less aggressive, committing fewer fouls compared to weekend games.

Research on European football participation provides further insight into the impact of playing in multiple competitions on league performance. Poli, Besson and Ravenel (2015) indicated that involvement in the Champions League or Europa League does not significantly affect domestic league results. In contrast, Moffat (2020) found that smaller league teams benefited domestically from Europa League participation, likely due to increased resources and experience gained. Furthermore, Cabras, Delogu and Tena (2022) used deep learning models to assess the effects of Champions League matches on domestic league performance. The findings revealed that smaller teams often performed worse in their league games following European matches, while larger teams showed more resilience to fixture congestion. This research guides our analysis to test whether similar effects are seen in smaller teams when considering both their total market value and squad size.

Despite the growing body of research on fixture congestion and performance, significant limitations impact the generalizability of these findings. Many studies, such as those by Carling and Dupont (2012), Dellal and Orhant (2015), and Dupont and Wisløff (2010), focus on data from a single team, which restricts the applicability of the results to other teams or leagues with different tactical styles, playing strategies, and fitness levels. These studies often involve a limited number of matches or players, reducing their statistical robustness. For instance, Carling and Dupont (2012) analyzed just eight matches over 26 days, which may not fully capture the long-term effects of fixture congestion over an entire season or across various competition formats.

The focus on individual performance metrics, such as distance covered or touches, as emphasized by Dellal and Orhant (2015), overlooks the broader tactical and psychological impacts of fixture congestion on team performance. Additionally, external factors such as travel distance, which can significantly affect team performance after midweek games, are frequently neglected. These factors are particularly relevant for teams that travel long distances between fixtures but are rarely examined in existing studies.

Another overlooked aspect is the importance of the tournament round. Later-stage cup matches are often more demanding and may pose greater recovery challenges, especially for larger teams expected to progress further in tournaments. Understanding how the importance of the round affects recovery could provide deeper insights into fixture congestion's impact on performance.

The effect of fixture congestion on physical performance and injury risk has also been studied. Carling and Dupont (2012) showed that, while total distance covered by players varied, high-intensity running performance remained stable, with similar injury rates during congested and non-congested periods. However, the recovery time for injuries was shorter during congested periods. Similarly, Dupont and Wisløff (2010) found that players who participated in two matches per week faced a significantly higher risk of injury compared to those who played one match, even though physical performance metrics such as total distance and high-intensity running were not greatly affected.

Many studies, such as Scoppa (2015), primarily focus on international tournaments like the FIFA World Cup or UEFA European Championship, limiting their relevance to domestic leagues where congestion dynamics, travel demands, and recovery routines differ. Cabras et al. (2022) similarly focused on Champions League participation, narrowing the analysis to elite teams and ignoring the experiences of smaller teams in lower divisions. This limits the understanding of how fixture congestion impacts less prominent teams with fewer resources and smaller squads. While deep learning models used in such analyses provide powerful insights, they often function as "black boxes" Ribeiro, Singh and Guestrin (2016), offering limited interpretability. This lack of transparency makes it difficult to pinpoint which specific factors, such as player fatigue, match location, or opponent strength, are most influential in performance declines after congested periods.

This study investigates how domestic cup matches affect a team's league performance in two ways. First, we examine whether winning or losing a cup match influences the outcome of their next league game, focussing on how the emotional boost or setback might impact performance. Second, we analyze whether participating in a cup match, regardless of the result, affects league performance compared to teams that did not face the added charge of an extra mid-week cup fixture. By looking at both momentum from a win and participation, we aim to understand how cup participation impacts the performance of a team in their next league fixture.

The data for this study come from detailed records of domestic cup and league matches in four major European football nations: England, Germany, the Netherlands, and Portugal, covering season 2012-2013 till season 2023-2024. We collect information on team performance in cup fixtures, league standings, and key team characteristics such as market value, squad composition, and age. In addition, travel distances and recovery periods between cup and league matches are constructed to control for factors that may affect league performance. This dataset allows us to analyze the effects of both winning and participating in cup competitions on subsequent league outcomes.

For estimating the causal effect of winning or participating in a domestic cup match on subsequent league performance, we leverage the random assignment of cup fixtures. The drawing process pairs teams with opponents of varying strength, creating natural variation that influences match outcomes without bias. Winning a cup match is expected to positively impact a team's momentum, while losing might worsen their performance in the next league fixture. For teams not participating in the cup round, the effect is expected to be different, possibly due to an increase in rest or preparation time. We use two complementary empirical approaches to analyze these dynamics.

Our first empirical approach uses the inherent randomness in cup draws by using two instrumental variables: the opponent's position in the league from the previous season and the division they played in. These variables are closely related to the probability of winning a cup fixture but are assumed to be exogenous with respect to performance in the next league fixture, making them suitable for isolating the causal effect of winning. To strengthen the argument for exogeneity, we validate our instruments by performing tests for independence between fixture assignments and opponent postition and division, confirming that they are assigned independently of potential unobserved confounders. This validation, combined with official cup regulations that require random draws, supports the independence and exclusion assumptions necessary for a robust analysis.

The second empirical approach focuses on the effect of participating in a cup round. Participation can cause physical and mental fatigue, which can affect performance in the following league fixture. By comparing the league outcomes of teams that participate in a cup fixture to those that do not, we capture a broader impact of fixture congestion. We include a set of control variables like travel distance, number of days between matches, and team characteristics such as squad depth and market value to mitigate concerns about omitted variable bias and ensure robustness in our findings.

The findings of this research show that, in general, earlier cup rounds, winning a cup fixture significantly improves performance in the next league fixture. We show that especially for lower market value teams (from lower divisions), winning a cup fixture improves performance in subsequent league fixtures. However, for teams with higher market value, the effect is less strong and statistically insignificant. Furthermore, participation in a cup match, regardless of the outcome, does not show a significant effect on the performance of the league in the different rounds. These findings suggest that smaller teams may benefit more from cup success in terms of league momentum, while larger teams with greater resources are less affected by fixture congestion.

Although there is a lot of research on the physical and psychological toll of fixture congestion, particularly with regard to the increased risk of injury and the reduced recovery time (L. Pillay, 2022; FIFPro, 2020), less attention has been paid to its impact on team performance after cup fixtures. Most of the literature focuses on the effect of international fixtures such as the UEFA Champions League and the Europa League on league performance. The literature lacks the effect of domestic cup matches on league performance. This study seeks to fill this gap by investigating how winning a domestic cup match affects a team's performance in their subsequent league fixture, using advanced econometric methods to account for potential biases in the data. The results of this research are interesting not only for teams participating at the highest level but also for football teams in lower divisions in European football pyramids.

The remainder of this paper is organized as follows. In Chapter 2, we describe the background of domestic cup competitions, the data collection and do preliminary data analysis. In Chapter 3, we explain our empirical strategy. In Chapter 4, we show the main estimation results. In Chapter 5, we test the robustness of our findings and extend our analysis with heterogeneity analysis. In Chapter 6, we conclude and discuss the results of the research.

Chapter 2

Background and data

2.1 Background

Football is the most popular sport in Europe, and domestic cup competitions are an important part of national football cultures. Tournaments like the FA Cup in England, DFB Pokal in Germany, KNVB Beker in the Netherlands, and Taça de Portugal in Portugal bring an extra level of excitement to the football season. These knockout-style tournaments draw teams from all levels of football, creating matchups between top-tier clubs and lower-league sides, with the potential for thrilling underdog stories.

The structure of these cup competitions increases their unpredictability. In contrast to league formats that reward consistent performance over a season, cup fixture can be determined by a single game, where the luck of the draw can play a crucial role. Small clubs, often lower-league or even semi-professional teams, can face football giants in early rounds, leading to "David vs. Goliath" scenarios. An iconic example is Wigan Athletic's 2013 FA Cup win, where, despite battling relegation in the Premier League, they defeated the top club Manchester City 1-0 in the final. Similarly, in 2023 DFB Pokal, third-division side 1. FC Saarbrücken shocked the football world by reaching the semifinals, eliminating Bundesliga sides such as SC Freiburg along the way.

Each competition has its own unique format. The FA Cup consists of 12 rounds, with lower- and non-league teams competing in the first six qualifying rounds, while Premier League and Championship sides join in the third round. The DFB Pokal follows a six-round structure, with all Bundesliga and 2. Bundesliga teams entering in the first round, competing alongside lower-tier clubs and regional cup winners. In the Netherlands, the KNVB Beker features eight rounds, with Eredivisie teams entering in the third round, and lower-tier teams advancing from earlier stages. The Taça de Portugal also has eight rounds, with the teams from the Primeira Liga joining in the third round.

These competitions are scheduled alongside regular league fixtures, often leading to fixture congestion for teams balancing both league and cup ambitions, and sometimes even international cup competitions. Fixtures scheduled during the week can affect squad resources, leading to fatigue, injuries, and inconsistent performance. Even with these challenges, cup competitions are still fan favorites because they bring out the excitement and unpredictability of football. The possibility of suprising upsets, excitements of underdog wins, and the high stakes of the knockout rounds contribute to the attractiveness of these competitions in the entire football pyramid of an association.

Table 2.1: Standardized Definitions of Rounds in Cup Competitions

Round	**Stage**
1	Final
2	Semi-Finals
3	Quarter-Finals
4	Round of 16
5	Round of 32
6	Round of 64

2.2 Data

Data for this study were collected from multiple sources, providing detailed information on cup fixtures, league fixtures, league standings, and team characteristics in four major European football pyramids: England, Germany, the Netherlands, and Portugal. The reason why the cup competitions of these countries are included is that the draw of their cup competition is entirely random, whereas cup draws of other top football nations such as France, Italy and Spain depend on seedings to protect larger clubs from facing each other until later rounds. The data set spans multiple seasons from 2012 to 2023 and is composed of three main components: cup fixtures, league fixtures, league standings, and team characteristics. Data on league standings and fixtures were sourced from API-Football, a database that provides detailed football statistics, including fixture details, fixture outcomes, and locations (API-Football, 2024). From these variables, we can calculate recovery times between fixtures and travel distances for the cup fixtures. Team financial data and squad decomposition were sourced from Transfermarkt, a platform for financial data and squad decomposition. Franck and Nüesch (2012) state that player valuations are determined by industry experts, who consider factors such as salaries, signing fees, bonuses, and transfer fees.

Before analyzing the data, we implemented several pre-processing steps to ensure consistency. First, we included only cup fixtures where the interval between the cup fixture and the following league fixture was five days or fewer. We do this to ensure that we only capture the immediate impact of a cup fixture on the subsequent league fixture. There could be interference with fixtures in international European competitions or other leagues if the interval would be longer than five days. Furthermore, league fixtures from the Netherlands and Portugal prior to 2016 were excluded due to incomplete data on league standings and fixtures for that period, as the data provider (API-Football) only offers information starting from 2016. Data prior to 2012 is not available from this provider, and while other providers may have such data, it is generally not accessible without cost or through free plans. Although the data from API-Football is well-documented, comprehensive access requires a paid subscription. Additionally, data on non-league (semi-professional) teams in England and Germany were unavailable. These non-league teams participate in competitions outside the top four English leagues, the top three German leagues, and the top two leagues in the Netherlands and Portugal. As such, data for these amateur football teams were not included in this study due to the difficulty in obtaining detailed information. Financial characteristics of football teams in the higher divisions were generally complete; however, some information was occasionally missing for teams in the lower leagues. To address these gaps, a Min-Max imputation technique was applied, using the lowest 5% of the distribution in the lowest available league to estimate missing values.

To allow cross-country comparisons of domestic cup competitions, we standardized the naming of the rounds across all tournaments. This standardization enables direct comparisons of equivalent stages, such as finals, semifinals, and quarterfinals. For instance, we refer to the final as Round 1, the semi-finals as Round 2, and the quarter-finals as Round 3. The standardized round definitions used in this study are described in Table 2.1. The schedules for the FA Cup, DFB Pokal, KNVB Beker, and Taça de Portugal show a relatively consistent structure, facilitating cross-competition comparison. For example, the Final (Round 1) is consistently scheduled in May for all four competitions, marking the end of the domestic football season. Similarly, Semi-Finals (Round 2) are typically played in April, and Quarter-Finals (Round 3) occur between January and March. These timing similarities make it feasible to compare the same stages of cup competitions across different countries. Although some countries have more rounds in their domestic cup competitions, the earliest stages often involve semi-professional or amateur teams for which detailed league fixture information is not available, as previously discussed. Therefore, we include only the six main rounds of the domestic cup competitions in our analysis to ensure data consistency and comparability.

This paragraph introduces the variables used in the analysis, with detailed definitions provided in Appendix A. The primary outcome, League Performance, measures the points a team earns in its first league fixture following a cup match (three points for a win, one for a draw, zero for a loss). Cup Participation is a binary variable indicating if a team played in a cup round, while Cup Win shows if the team won the cup fixture.

The intensity of a cup fixture is captured by Extra Time, indicating if the cup match required extra time. Team Position and Opponent Position represent league rankings of team and opponent in the previous year, while Recovery Days tracks the interval between the cup and the subsequent league fixture. Distance measures the travel distance for the away team, derived from team locations.

Team characteristics include Team Size (number of squad members), which is a proxy for squad depth. Foreign Players counts the number of non-domestic players, adding potential diversity and experience. Mean Age reflects squad maturity, while Mean Market Value serves as a proxy for team skill, and Total Market Value indicates overall financial strength. Lastly, Country Code, based on UEFA Country Coefficients (UEFA, 2025), reflects the relative strength of football associations.

After filtering the data to include only league fixtures that occurred within five days of a domestic cup fixture, we retained around 30% of the original observations. This ensures that we analyze the immediate impact of cup participation on league performance, avoiding confounding effects from extended recovery periods or other fixtures.

Table 2.2: Summary of Fixtures and Teams per Round by Country/Cup

Country	**Cup**	**Round**	**Years**	**Teams (Mean)**	**Fixtures**	
					Mean	**Total**
England	FA Cup	1	2012-2023	2.0	1.0	12
		2		3.9	2.0	24
		3		7.8	3.9	47
		4		15.4	8.0	96
		5		30.9	16.0	192
		6		62.2	32.0	384
		7		38.7	20.1	241
		8		76.8	40.0	480
Germany	DFB Pokal	1	2012-2023	2.0	1.0	12
		2		3.8	2.0	24
		3		7.8	4.0	48
		4		15.6	8.0	96
Netherlands	KNVB Beker	1	2016-2023	2.0	1.0	7
		2		4.0	2.0	16
		3		8.0	4.0	32
		4		16.0	8.0	64
		5		29.8	14.9	119
		6		53.6	27.0	216
Portugal	Taça de Portugal	1	2016-2023	2.0	1.0	8
		2		4.0	3.0	24
		3		7.6	4.0	32
		4		29.8	15.4	123
		5		55.5	29.5	236
		6		79.9	42.6	341
		7		48.8	48.3	386
Total						**3260**

Cup Participation

To estimate the effect of cup participation on subsequent league performance, we construct the variable Cup Participation to indicate whether a team plays in a given cup round r, regardless of the fixture outcome. This variable serves as the basis for comparing league performances between teams that participated in a cup fixture and those that did not. Specifically, we evaluate League Performance $r + 1$, defined as the points obtained in the league fixture within five days after the cup round $r + 1$.

For teams that lose in cup round r and thus do not progress to round $r+1$, we assign their next league fixture within the five-day window as their "League Performance $r + 1$". This approach ensures a consistent comparison by linking every team, whether they advance in the cup or not, to a subsequent league performance. Importantly, for non-participating teams, we confirm that no other fixtures occur within the five-day period before the evaluated league fixture, ensuring that the results reflect the direct effect of cup participation and are not confounded by other fixtures or potential fatigue.

2.3 Exploratory Data Analysis

Table 2.3 shows key variables, split by Cup Participation and Cup Win for a cup round. Teams that participated in cup competitions generally have higher Total Market Values, particularly in the low and medium categories, compared to non-participants, with a significant difference in the lower value category ($p < 0.0001$). Similarly, teams that win a cup fixture tend to have higher values in both the low and high categories, with marginal significance in the high category ($p = 0.05$), suggesting that teams with higher market value are more likely to succeed. Team Size shows little variation between participants and non-participants in the low and medium groups. However, high-value teams that participate in cup competitions have significantly larger squads ($p = 0.03$), indicating that larger squads help manage the added fixture load. The percentage of Foreign Players is also higher for participants in the low category ($p = 0.01$), implying that more diverse squads are likelier to progress to later rounds of the cup competition. The Recovery Days variable shows a significant difference ($p < 0.0001$), with cup participants having slightly more rest between fixtures.

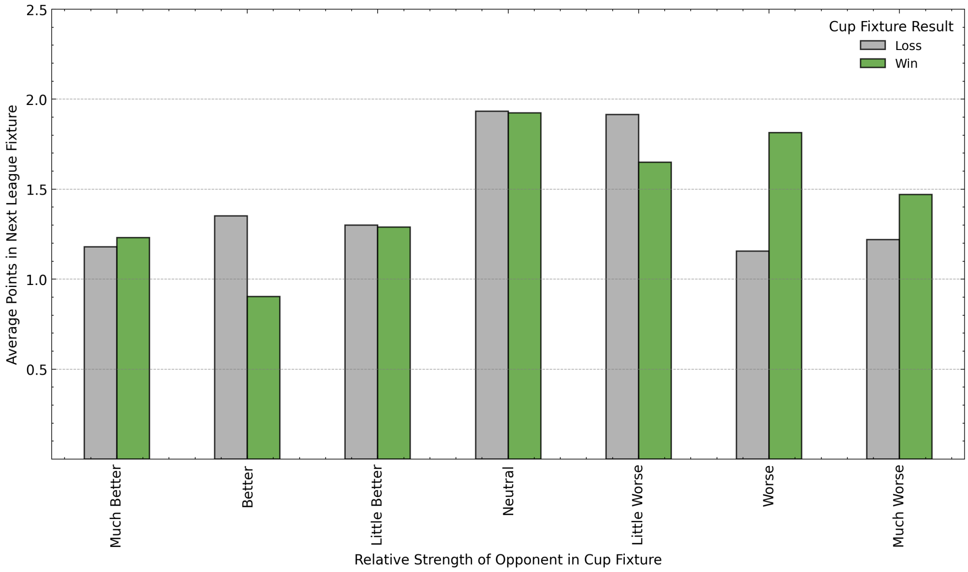
The correlation heatmap in Appendix B.1 provides valuable insights relations between key variables of our research. League Performance in the fixture immediately following a cup fixture shows a moderate positive correlation with Mean Market Value and Total Market Value, indicating that financially stronger teams tend to perform better in league fixtures after a cup round. This relationship weakens slightly in the second league fixture, suggesting a potential decrease in the financial advantage over time. Cup Win also shows a weak positive correlation with financial indicators, implying that wealthier teams are more likely to succeed in cup fixtures, potentially impacting subsequent league performance.

Variables such as Extra Time and Distance show low correlations with League Performance, indicating that, at first impression, fatigue from extra time or travel appears to have minimal impact on league outcomes. Additionally, Opponent Position and Division have low correlation with league results, providing context on the competition level without confounding league performance. The positive correlation between League Performance r and League Performance r+1 highlights consistency in team form over multiple league fixtures, probably reflecting strength and depth of a team.

Moreover, Team Size shows a small positive correlation with League Performance r, suggesting that larger squads may have a small advantage. In contrast, Mean Age has a small negative correlation, indicating that older squads might perform slightly worse in league fixtures following cup fixtures. Lastly, Country and Team Position have moderate negative correlations with League Performance, reflecting potential differences in competitive intensity and team quality across leagues.

Figure 2.1 shows average points earned in the league fixture immediately following a cup fixture, categorized by the relative strength of the opponent. Categories are determined by the difference between Opponent National Position and Team National Position, with "Much Better" indicating a higher-ranked opponent and "Much Worse" a lower-ranked opponent. Teams generally perform better in the next league fixture after winning a cup fixture. However, teams that defeat a much stronger opponent tend to earn fewer points in the subsequent league fixture, possibly due to fatigue. In contrast, teams that win against an opponent of similar or weaker

Figure 2.1: Average Points in next League Fixture by Opponent Strength.

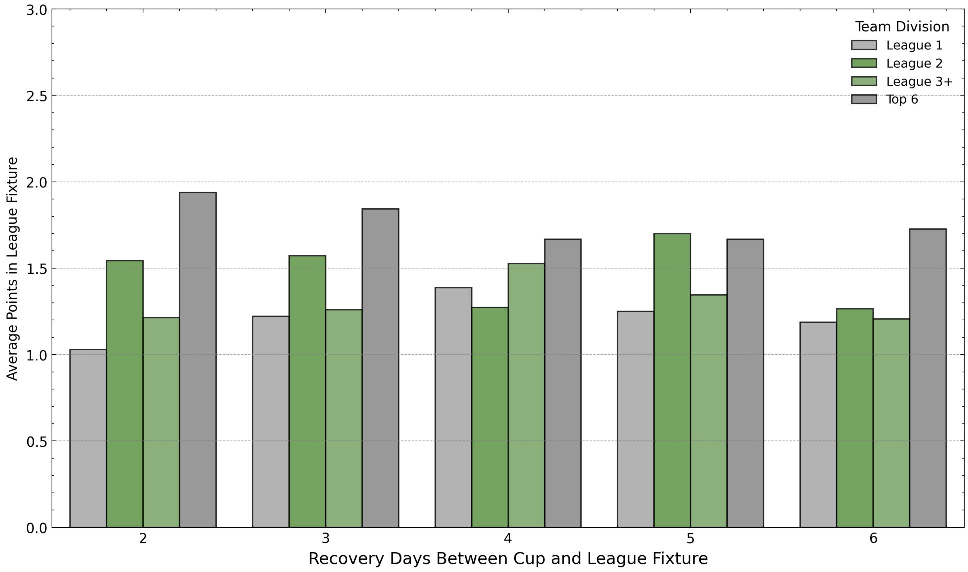


Note: This figure presents the average points obtained in the league fixture following a cup fixture, categorized by the relative strength of the cup opponent. Categories reflect the difference between Opponent National Position and Team National Position: "Much Better" and "Much Worse" refer to opponents from much higher or lower leagues, respectively (with more than 20 positions difference). Smaller differences indicate teams within the same league, with "Little Better" and "Little Worse" referring to teams within a 5-position difference.

standing often show improved league performance in the following fixture. These observations suggest that both the cup fixture outcome and the relative strength of the opponent could play a role in influencing league performance, providing a foundation for further analysis.

Figure 2.2 shows the average points in the league fixture following a cup fixture, categorized by recovery days and team division. Surprisingly, teams in the "Top 6" of Division 1 tend to maintain higher average points with fewer recovery days (2–3 days), which may reflect their team strength and depth, allowing them to perform well even with limited rest. This could also be attributed to facing relatively weaker league opponents due to scheduling by football associations. In contrast, other divisions, particularly League 2 and 3+, show a trend of slightly higher average points as recovery days increase, suggesting these teams benefit more from additional rest days to increase league performance.

Figure 2.2: Average Points in the next League Fixture by Recovery Days and Team Division.



Note: This figure shows the average points obtained in the league fixture following a cup fixture, categorized by recovery days and team division. Categories represent different team divisions, with "Top 6" referring to the top six teams in Division 1, and separate groups for teams in Division 1, Division 2, and Division 3 or lower. Recovery days indicate the time between cup and league fixtures, illustrating the effect of rest periods on subsequent league performance.

Table 2.3: Summary Statistics by Participation and Cup Win

	Participation		p-value	Cup Win		p-value
	Yes	No	F-test	Win	No Win	F-test
Total Value						
Low	115.14	86.17	< 0.0001	115.96	86.39	< 0.0001
	(5.37)	(4.81)		(5.38)	(5.15)	
Medium	550.24	536.42	0.61	533.09	540.20	0.61
	(12.23)	(24.15)		(10.73)	(20.60)	
High	929.83	969.79	0.40	878.25	974.26	0.05
	(21.99)	(40.82)		(27.67)	(35.83)	
Mean Value						
Low	3.04	2.26	< 0.0001	3.33	2.32	< 0.0001
	(0.14)	(0.12)		(0.15)	(0.14)	
Medium	14.63	13.92	0.27	15.61	15.20	0.27
	(0.32)	(0.55)		(0.34)	(0.65)	
High	23.17	25.16	0.13	24.16	26.54	0.13
	(0.45)	(1.12)		(0.59)	(1.58)	
Team Size						
Low	31.92	31.69	0.30	31.78	31.74	0.30
	(0.14)	(0.17)		(0.14)	(0.18)	
Medium	38.98	38.87	0.61	38.95	38.86	0.61
	(0.14)	(0.17)		(0.15)	(0.18)	
High	50.23	48.10	0.03	49.32	49.54	0.03
	(0.68)	(0.65)		(0.49)	(1.02)	
Foreigners						
Low	12.49	11.63	0.01	12.52	11.52	0.01
	(0.21)	(0.24)		(0.20)	(0.26)	
Medium	20.82	20.51	0.21	20.55	20.62	0.21
	(0.16)	(0.18)		(0.16)	(0.20)	
High	31.31	30.74	0.43	31.46	30.79	0.43
	(0.43)	(0.58)		(0.47)	(0.56)	
Mean Age						
Low	23.62	23.57	0.31	23.48	23.44	0.31
	(0.03)	(0.04)		(0.03)	(0.04)	
Medium	25.28	25.28	0.94	25.11	25.12	0.94
	(0.03)	(0.04)		(0.03)	(0.04)	
High	27.40	27.30	0.46	27.12	27.06	0.46
	(0.11)	(0.07)		(0.09)	(0.09)	
Distance						
Low	55.77	55.98	0.96	56.74	56.22	0.90
	(2.65)	(2.75)		(2.44)	(3.22)	
Medium	156.59	156.44	0.98	144.20	144.67	0.92
	(3.33)	(3.54)		(2.71)	(3.43)	
High	416.40	381.40	0.30	332.66	348.86	0.54
	(24.48)	(22.78)		(14.63)	(22.23)	
Recovery Days	3.13	2.73	< 0.0001	3.15	3.04	< 0.0001
	(0.03)	(0.05)		(0.03)	(0.04)	
Extra Time	0.13	0.14	0.52	0.14	0.16	0.52
	(0.01)	(0.02)		(0.01)	(0.02)	
Outcome Variable:						
Points in Next Fixture	1.57	1.20	–	1.53	1.34	–
	(0.05)	(0.06)		(0.05)	(0.06)	
Observations	620	497		609	429	

Chapter 3

Empirical strategy

This paper estimates the effect of participating and winning a domestic cup on subsequent league performance in European football, by leveraging the randomness of cup draws as an exogenous source of variation. The results of the draws create natural variations in the strength of the opponent, influencing the likelihood of winning and progressing in the tournament. The variation in opponent strength from random draws has been used in prior research to analyze its impact on financial outcomes and the number of goals scored (Simmons & Walker, 2010; Wunderlich, Seck & Memmert, 2023). There is evidence that facing a stronger opponent in a cup competition is associated with an increased likelihood of early elimination, while drawing a weaker opponent is correlated with progressing further in a cup tournament (Dagaev & Suzdaltsev, 2018; Cea, Durán, Guajardo, Siebert & Zamorano, 2020).

To frame the methodology, we start by describing the potential outcome framework, a tool for causal inference, as first introduced by Neyman (1923) and formalized by Rubin (1974). In this framework, we represent the treatment effect of domestic cup participation using potential outcomes. Let Y_i be the league performance in terms of points in the next league fixture of team i, and let $W_i \in \{0, 1\}$ denote the binary treatment variable, where $W_i = 1$ indicates participation (win) in the cup, and $W_i = 0$ otherwise. The potential outcomes $Y_i(1)$ and $Y_i(0)$ represent the team's league performance conditional on participation and non-participation, respectively. The outcome realized Y_i is shown in Equation 3.1:

$$Y_i = W_i Y_i(1) + (1 - W_i) Y_i(0) \tag{3.1}$$

Equation 3.1 highlights the fundamental problem of causal inference, as only one potential outcome is observed for each team (Holland, 1986). The individual treatment effect τ_i is defined in Equation 3.2:

$$\tau_i = Y_i(1) - Y_i(0) \tag{3.2}$$

However, τ_i cannot be directly calculated. Instead, we estimate the average treatment effect (ATE), as shown in Equation 3.3:

$$\tau_{ATE} = E[Y_i(1) - Y_i(0)] \tag{3.3}$$

To make this estimation possible, the framework relies on several assumptions. The first assumption is the Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1980). SUTVA assumes that the potential outcomes for each unit (team) depend only on its own treatment status, and there is no interference between units nor hidden variations in the treatment itself. In this study, it implies that one team's participation in the cup does not directly affect the performance of other teams. The second assumption is Ignorability, or Unconfoundedness, which assumes that, after controlling for observed covariates X_i, the assignment of the treatment W_i is as good as random. This is formally written in Equation 3.4:

$$(Y_i(1), Y_i(0)) \perp W_i \mid X_i \tag{3.4}$$

This assumption ensures that any differences in league performance between participating and non-participating teams are due to cup participation itself, rather than unobserved confounders. The third assumption is Common Support, which requires that for every team, there is a positive probability of participating and not participating in the cup, as shown in Equation 3.5:

$$0 < P(W_i = 1 \mid X_i) < 1 \quad \forall X_i \tag{3.5}$$

While these assumptions are essential for identifying causal effects, in this study, the Ignorability assumption in Equation 3.4 does not hold. Cup participation in England, Germany, the Netherlands, and Portugal is not randomly assigned, as it depends on team quality and other unobserved factors. Stronger teams are more likely to progress in cup competitions and perform better in the league, leading to selection bias and making OLS estimates unreliable and biased (Angrist, Imbens & Rubin, 1996; Heckman & Vytlacil, 2005). This issue, often referred to as endogeneity, arises because factors that influence cup participation are also likely to affect league performance. To address this, we use Instrumental Variables (IV), which allow us to isolate the exogenous variation created by the randomness of the cup draw and estimate the causal effect of cup participation on league performance. IV methods, introduced by Imbens and Angrist (1994), are widely applied in settings where endogeneity is a concern, as they exploit external sources of variation (Imbens, 2014; Wooldridge, 2010). They depend on the availability of instruments; a variable that affects the treatment intake but does not directly affect potential outcomes.

To visualize this framework, we can represent the relationships between the treatment W, the outcome Y, the instrument Z, and the unobserved confounder U using a Directed Acyclic Graph (DAG), as shown in Figure 3.1 from (Abadie & Cattaneo, 2018). This DAG illustrates how the instrument Z influences the treatment W, which in turn affects the outcome Y, while the unobservable confounder U directly affects both W and Y, making conditioning on U unfeasible.

In our setting, the strength of the team's cup opponent serves as the instrumental variable Z. The instrument influences whether the team progresses in the cup but is unrelated to unobserved factors that could affect subsequent league performance.

The use of IV allows us to estimate the Local Average Treatment Effect (LATE), as introduced by Imbens (2014). LATE captures the causal effect of cup participation for teams whose

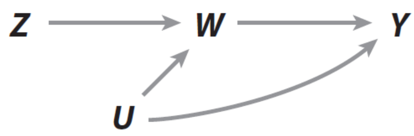


Figure 3.1: Instrumental Variable in a Directed Acyclic Graph. The graph shows the treatment, W; an unobservable confounder, U; the outcome variable, Y; and an instrument, Z, affecting the treatment.

participation is influenced by the instrument—referred to as "compliers." In our study, these are teams whose participation in the cup depends on the strength of their opponent. LATE is particularly helpful here because it estimates the treatment effect specifically for teams influenced by the randomness of the draw, offering insights into this specific group rather than the overall population, which would be captured by the Average Treatment Effect (ATE).

Formally, let Z denote the instrument, which in our case is the strength of the opponent determined by the random cup draw. The treatment W, cup participation or win, is determined by both the instrument Z and the team's underlying characteristics. The potential outcomes for cup participation can be represented as W_1 if the team faces a weak opponent and W_0 if they face a strong opponent. The observed treatment status W is given by Equation 3.6:

$$W = ZW_1 + (1 - Z)W_0 \tag{3.6}$$

where W_1 represents the potential cup participation (win) under assignment to a weak opponent, and W_0 represents the potential cup participation (win) under assignment to a strong opponent. This framework accounts for the fact that not all teams comply with their treatment assignment. For instance, some teams may win regardless of opponent strength (always-takers), while others may lose regardless (never-takers).

If we assume that the treatment effect (i.e., the impact of winning or advancing in the cup) does not vary across teams, we can use the Wald estimator to estimate the LATE. The instrumental variable (IV) estimator can be represented as shown in Equation 3.7:

$$\tau_{IV} = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)} = \frac{E[Y_1 - Y_0 \mid Z = 1] - E[Y_1 - Y_0 \mid Z = 0]}{E[W_1 - W_0 \mid Z = 1] - E[W_1 - W_0 \mid Z = 0]} \tag{3.7}$$

This IV estimator isolates the exogenous variation in the treatment caused by the random assignment of opponents. The numerator reflects the variation in league performance caused by the instrument, while the denominator captures how the instrument affects the probability of winning a cup fixture.

In the presence of heterogeneous treatment effects, where the impact of winning a cup fixture might vary between teams, the Wald estimator may fail to capture the average effect accurately. This happens because teams vary in how they respond to the treatment, with the instrument influencing some teams to win (compliers) while leaving others unaffected (always-takers or never-takers). A solution is to investigate if the effect of cup participation differs significantly across teams from different divisions (i.e., lower vs. higher divisions). We can do so by splitting the sample based on team characteristics.

As shown in Equation 3.8:

$$\tau_{IV} = \frac{E[Y_1 - Y_0 \mid W_1 > W_0]}{E[W_1 - W_0]} \tag{3.8}$$

This expression reveals that τ_{IV}, the IV estimate, equals τ_{LATE} under the assumption of monotonicity. Monotonicity implies that no team defies the treatment assignment; for example, facing a stronger opponent does not increase the chance of progressing in the cup. In our context, it means that weaker opponents increase the probability of advancing in the cup.

By using instrumental variables and the LATE framework, we can isolate the causal effect of domestic cup participation on league performance for compliers—teams whose cup participation is influenced by the random assignment of opponents. As discussed before, we will test two empirical applications: the first is the causal effect of participating in the domestic cup, the second is the causal effect of winning a fixture in the domestic cup on league performance. Although we discussed the methodology above for one instrumental variable for simplicity, in the empirical studies we will use two instrumental variables to capture both opponent position and the division they play in. The first stage of the 2SLS approach models the endogenous treatment variable, which in our case is either participation in or winning a cup fixture, as a function of the instruments and additional covariates.

Formally, the first-stage regression can be written as shown in Equation 3.9:

$$W_{i,r} = \pi_{0,r} + \pi_{1,r} Z_{1,i,r} + \pi_{2,r} Z_{2,i,r} + \gamma X_{i,r} + \epsilon_{i,r} \tag{3.9}$$

where $W_{i,r}$ is the binary variable indicating cup participation or win for team i in round r, $Z_{1,i,r}$ represents the opponent's division, and $Z_{2,i,r}$ represents the opponent's position in the previous season. The term $X_{i,r}$ denotes a vector of covariates that includes team characteristics such as team strength, home advantage, and prior league performance, while $\epsilon_{i,r}$ is the error term. This first stage predicts cup participation as a function of the exogenously assigned opponent strength, captured by the instruments.

In the second stage of 2SLS, the predicted values from the first stage ($\hat{W_{i,r}}$) are used in place of the actual treatment variable to estimate the causal effect on league performance $Y_{i,r}$, as shown in Equation 3.10:

$$Y_{i,r} = \beta_{0,r} + \beta_{1,r} \hat{W_{i,r}} + \delta X_{i,r} + \eta_{i,r} \tag{3.10}$$

where Y_i represents the league performance outcome, $\hat{W_{i,r}}$ is the fitted value from the first stage that represents the exogenously adjusted cup participation, X_i remains the set of covariates, and $\eta_{i,r}$ is the error term. This second stage captures the causal relationship between cup participation and subsequent league performance by leveraging the exogenous variation provided by the instruments.

The 2SLS estimator for the effect of cup participation on league performance can be written as in Equation 3.11:

$$\hat{\beta}_{2SLS} = \left(Z'X\right)^{-1} Z'Y \tag{3.11}$$

where Z includes both instruments, the opponent's division and rank, X represents exogenous covariates, and Y is the league performance outcome. This estimator provides a weighted average of local average treatment effects (LATE) for teams whose cup participation is influenced by the instruments. Importantly, the 2SLS estimator captures the causal effect for the subset of teams (compliers) whose participation or win is driven by the exogenous variation in the instruments.

In Table 3.1, the model specifications with different sets of control variables are listed. Definitions for these control variables can be found in Appendix A. We will test the causal effects using seven model specifications in a stepwise manner to ensure the robustness of control variables through incremental inclusion. The equations for each model specification are provided in Appendix C. We begin with the baseline model (Model 1), which captures the direct effect of winning the cup fixture without any control variables. This corresponds to the model definitions in Equations 3.9 and 3.10 without the control variables $X_{i,r}$. Model 2 introduces Team Position to control for baseline team strength. Model 3 includes Distance to account for potential fatigue from travel, while Model 4 adds Days till League Fixture to reflect recovery time. Model 5 incorporates Extra Time as a control for fixture intensity of the cup fixture. In Model 6, team-specific characteristics such as Team Size, Total Market Value, and Mean Age are included, based on exploratory data analysis (EDA) findings, while avoiding multicollinearity (e.g., Mean Market Value and Foreign Players were excluded due to their high correlation with Total Market Value). Finally, Model 7 adds Country Code to adjust for cross-country differences in competitiveness and cup importance.

Table 3.1: Model Specifications

Model	**Group (Added)**	**Controls (Added)**
1	-	-
2	Base Strength	Team Position
3	Distance	Distance
4	Recovery	Days till League Fixture
5	Intensity	Extra Time
6	Team Characteristics	Team Size, Total Market Value, Mean Age
7	Country	Country Code

Validation of IV assumptions

This section outlines the assumptions underlying the instrumental variable (IV) approach to estimate the causal effect of domestic cup participation and winning on subsequent league performance. Following the framework of (Mogstad, Torgovitsky & Walters, 2021), we examine the assumptions of independence, exclusion, relevance, and monotonicity when using multiple instruments, ensuring that opponent position and division provide exogenous variation for unbiased 2SLS estimation.

The independence assumption requires that the instruments are uncorrelated with any unobserved confounders affecting league performance. This assumption is crucial because, as any correlation between the instruments and unobserved factors would violate exogeneity, makeing the causal estimates invalid. We leverage the inherent randomness in the domestic cup competition draws as the foundation for this assumption. For example, the FA Cup Rules for 2024-25 explicitly state that "The ties for each round of the Competition shall be determined by means of a draw" (The Football Association, 2024), ensuring that the assignment of opponents is randomized across position and division without seeding or systematic bias. This institutional rule ensures that opponent characteristics, such as division and league position, are exogenously determined with respect to the teams in the draw, thereby supporting the independence assumption.

To provide empirical support for independence, we perform chi-squared tests for independence between fixture assignments (Fixture ID) and opponent position and division across FA Cup rounds, following the methodological approach outlined by (Johnson & Wichern, 2018). The results, presented in Table 3.2, show that we fail to reject the null hypothesis of independence across all rounds, indicating that opponent position and division are randomly assigned. This consistency in randomness is further observed in similar tests conducted for domestic cup competitions in other countries (Appendix D), suggesting that the instruments are exogenous with respect to league performance outcomes. This evidence justifies their use as instruments in the IV analysis.

Table 3.2: Chi-2 Test for Independence by Round and Fixture Assignment in the FA Cup.

Round	**Fixture vs Division**			**Fixture vs Position**		
	Chi-2	**p-value**	**df**	**Chi-2**	**p-value**	**df**
1	11.48	.4041	11	84.00	.6009	88
2	44.06	.5537	46	334.88	.6420	345
3	85.96	.6578	92	886.17	.7832	920
4	248.02	.9284	282	1995.54	.3623	1974
5	544.10	.7485	567	4231.86	.8923	4347
6	1142.46	.4490	1137	8665.88	.6491	8717

Note: This table shows the Chi-2 test results for the independence of fixture assignments from opponent division and position in each FA Cup round. The Chi-2 statistic, p-value, and degrees of freedom (df) are reported. Failure to reject the null hypothesis supports the independence of fixture assignments, aligning with the random nature of the FA Cup draw and justifying these characteristics as instruments in IV analysis.

The exclusion restriction requires that the instruments affect subsequent league performance only through cup participation or winning a cup fixture, with no direct effect on league outcomes. This means that any influence of opponent characteristics on league performance should occur only via changes in cup participation status. Since the draw is random, opponent rank and division should theoretically have no other pathway to influence league performance except through participation.

To assess this, we examine whether any plausible alternative pathway exists for opponent position and division to impact the subsequent league performance of a team directly. In our analysis, as shown in Table 3.1, we estimate the effects across different model specifications with varying control variables. If we find stability across the different model specifications, we find indirect support for the exclusion restriction by suggesting that the effect operates primarily through cup participation or winning a cup fixture rather than any direct impact from the opponent position or division.

To further support the independence and exclusion assumptions, we perform robustness checks, detailed in Chapter 5. These include the use of an alternative instrument, travel distance, to test if the causal effect remains consistent. In addition, we perform a heterogeneity analysis to examine if the effects vary by subgroup, such as by team market value or squad size. These robustness tests help to find additional insights in the causal analysis and strengthen the IV assumptions.

The monotonicity assumption holds in expectation, as higher-ranked teams are generally more likely to progress in cup rounds, while lower-ranked teams face a higher probability of early elimination. However, occasional upsets mean this pattern is not absolute, as some lower-ranked teams do advance unexpectedly (Dagaev & Suzdaltsev, 2018; Cea et al., 2020). These rare deviations minimally impact the assumption's validity, supporting the use of monotonicity in our IV framework.

Likewise, the relevance assumption in our IV framework is supported by the strong relationship between the instruments (opponent position and division) and the likelihood of progressing in the cup. Higher-ranked teams are statistically more likely to advance against lower-ranked opponents, and this rank-based variation in opponent strength provides meaningful exogenous variation that directly influences the treatment. This relevance will be further validated if we observe statistically significant first-stage results with an F statistic of at least 10, indicating that opponent rank and division strongly predict progression or win probabilities across rounds. Meeting this condition ensures instrument relevance and confirms that our IV estimates capture the causal effect (Stock & Yogo, 2005).

Together, the official cup regulations, chi-squared tests for independence, and verifying stability across model specficiations we aim to provide strong support for the independence and exclusion assumptions. These theoretical and empirical analyses validate the use of opponent position and division as instruments in the IV analysis.

Chapter 4

Main results

This section presents the estimation results, focussing on the effect of participation or a win in a domestic cup round on the performance of the subsequent fixture in the league.

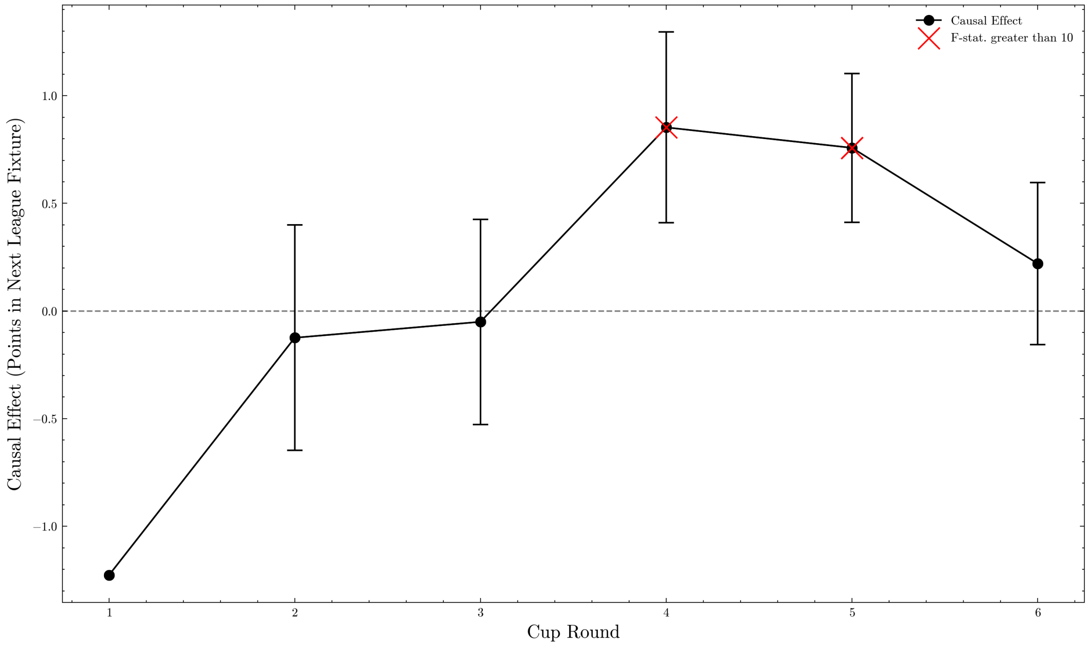
4.1 The effect of winning a cup round

When estimating the effect of winning a cup round on the points obtained in the subsequent league fixture right after the cup round. In this way, we compare the teams that won the cup round (treated teams) with the teams that did not win the cup round (non-treated teams). The goal is to measure whether teams that win the cup round experience a mental boost from the momentum of winning a cup fixture. In Figure 4.1 a summary of the causal effect per round of the cup can be seen, there are significant causal effects in earlier rounds of the cup competition. In round 1 (final) until round 3 (quarter final), no significant causal effect has been found.

The results presented in Table 4.1 measure the effect of winning a cup fixture in Round 1 (final) on subsequent league performance, using opponent position and division as instrumental variables (IV). In Panel A, the first-stage results highlight the relationship between the exogenous variables (opponent position and division) and the likelihood of winning the cup fixture. The coefficient on opponent division in Model 1 is positive and statistically significant at the 5% level, indicating that teams facing lower-division opponents are more likely to win the cup fixture. However, in subsequent models, the significance of this variable diminishes, suggesting potential issues with instrument strength in specifications with team and country characteristics in the control variables. The F-statistics, which range from 0.37 to 1.58, suggest that the instruments are weak in some specifications, especially when including more controls. In Panel B, the second-stage results demonstrate the effect of winning the cup fixture on league performance. The coefficients on the win variable vary across the models, switching between negative and positive values, with no specification yielding statistically significant results. The lack of statistical significance, combined with fluctuating effect sizes, raises questions about the robustness of the estimated effects. The R^2 values increase as more controls are added, indicating that additional covariates help to explain the variance in league performance, although the overall explanatory power remains limited in the simpler models, this can be mainly attributed to the small sample we use for the regression.

Moving to Table 4.2, which focuses on the effect of winning a cup fixture in Round 2 (semi-

Figure 4.1: Causal Effect of Winning a Domestic Cup Round on League Performance.



Note: This figure illustrates a summary of the estimated causal effects of winning a cup round on league performance, measured in points obtained in the subsequent league fixture. Each point represents the estimated effect for a specific cup round (1 through 6), shown with confidence intervals. The black line indicates the main effect, while the red crosses highlight rounds where the first-stage F-statistic exceeds 10, suggesting stronger instrument relevance.

final) on league performance, the first-stage results in Panel A show a consistently strong relationship between the instrumental variables and the likelihood of winning. The coefficient on opponent position is positive and highly significant across all models, with p-values below 0.01. Similarly, the opponent division variable is statistically significant at the 1% level, confirming the relevance of the instruments in predicting cup fixture outcomes. The F-statistics, which range from 8.52 to 15.84, often exceed the critical threshold, indicating that the instruments are strong. However, in panel B, the second stage estimates show no significant effect of winning the cup fixture on subsequent league performance. The coefficients on the win variable are consistently negative, but none reach statistical significance, suggesting that there is no clear impact of winning a cup fixture in Round 2 on league performance. The R^2 values remain relatively low, indicating that the model explains a small portion of the variance in league performance, despite the strong instruments in the first stage. Again we see that only 128 observations were available for estimating the effects, more than in round 1, but this sample size is still questionable for proper inference.

Table 4.3 presents the results for Round 3 (quarter-final). Panel A continues to show strong first-stage results, with opponent position and division both significantly predicting the likelihood of winning the cup fixture. The F-statistics remain high, above the threshold for weak instruments, confirming the relevance of the instruments. Also in this estimation we see that as we include more control variables in the model, especially for team characteristics, the F-statistics fall below the crititical threshold of 10, possibly suggesting a weak instrument. In panel B, the second stage results show no significant effect of winning the cup fixture on league

Table 4.1: Estimates of the effect of wining a cup fixture in round 1 (final) on subsequent league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = win cup fixture)							
Opponent Position	-0.03	-0.01	0.02	0.02	0.02	0.24	0.29
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.11)	(Inf)
Opponent Division	0.63**	0.40	-0.13	-0.03	-0.18	10.56	37.03
	(0.27)	(0.41)	(0.40)	(0.84)	(0.96)	(6.18)	(Inf)
Panel B: IV estimates (dependent variable = next league performance)							
Win Cup Fixture	-4.05	-3.13	3.28	3.56	3.43	-0.43	-1.24
	(3.97)	(10.05)	(4.82)	(5.88)	(6.31)	(2.39)	(Inf)
F-Statistic	0.37	0.45	2.14	1.35	0.96	1.58	1.00
R^2	0.04	0.11	0.52	0.52	0.55	0.93	1.00
Observations	10	10	10	10	10	10	10

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 1. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.2: Estimates of the effect of winning a cup fixture in round 2 (semi-final) on league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = win cup fixture)							
Opponent Position	0.03***	0.04***	0.03***	0.03***	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.27***	0.25***	0.25***	0.28***	0.28***	0.32***	0.30***
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Panel B: IV estimates (dependent variable = next league performance)							
Win Cup Fixture	-0.21	-0.15	-0.15	-0.12	-0.19	-0.20	-0.16
	(0.56)	(0.52)	(0.52)	(0.51)	(0.50)	(0.51)	(0.51)
F-Statistic	13.95	15.84	11.97	10.22	8.52	8.70	8.62
R^2	0.18	0.28	0.28	0.30	0.30	0.40	0.42
Observations	128	128	128	128	128	128	128

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 2. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

performance. Although the coefficients in the win variable are positive in some models, they do not reach statistical significance, and the R^2 values are relatively low, suggesting limited explanatory power.

Table 4.3: Estimates of the effect of winning a cup fixture in round 3 (quarter final) on league performance using random drawn opponent position and division as an IV

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Panel A: First stage (dependent variable = win cup fixture)							
Opponent Position	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.19***	0.16***	0.16***	0.16***	0.16***	0.17***	0.17***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Panel B: IV estimates (dependent variable = next league performance)							
Win Cup Fixture	0.20	0.06	0.09	0.07	0.07	-0.03	-0.03
	(0.51)	(0.53)	(0.54)	(0.54)	(0.54)	(0.50)	(0.50)
F-Statistic	15.96	16.94	13.01	10.39	8.61	8.50	8.20
R^2	0.14	0.20	0.21	0.21	0.21	0.28	0.30
Observations	202	202	202	202	202	202	202

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 3. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Table 4.4, which examines the effect of winning a cup fixture in Round 4 (Round of 16) on subsequent league performance, the first-stage results (Panel A) demonstrate a strong relationship between the instrumental variables and the likelihood of winning the cup fixture. Both variables are statistically significant at the 1% level across all models, with F-statistics ranging from 9.28 to 24.43, confirming the strength of the instruments.

In the second-stage results (Panel B), winning a cup fixture has a positive and statistically significant effect on league performance in Models 1 through 5, with coefficients ranging from 0.87 to 1.03 and p-values below 0.05. The inclusion of control variables in these models shows that team-specific characteristics and match-related factors slightly change the estimated effect. For example, when the team's previous position is included in Model 2, the coefficient on winning decreases, indicating that part of the win effect can be attributed to stronger teams being more likely to win. Similarly, controlling for distance traveled (Model 3) and recovery time (Model 4) reduces the magnitude of the effect, suggesting that fatigue and rest also play a role in determining league performance after a cup win.

As additional control variables related to team characteristics (such as squad size, average age, and financial indicators) are introduced in Model 6, the coefficient declines to 0.87 and loses statistical significance. This suggests that much of the observed performance boost from winning a cup fixture is driven by inherent team qualities, such as squad depth and financial strength, rather than the psychological momentum gained from the win itself. Therefore, we will investigate the differences in causal effects for teams with a larger squad size and a smaller squad size and teams with more financial resources and fewer financial resources in Section 5.2. In Model 7, the addition of country-specific controls does not substantially change the results, confirming that the win effect is not strongly influenced by broader league or country-specific factors.

The results for Round 4 suggest that winning a cup fixture can lead to improved league performance, but much of this effect is explained by team characteristics. As more controls are added, the magnitude of the win effect diminishes, suggesting that strong teams with more financial resources are less affected by winning a cup fixture.

Table 4.4: Estimates of the effect of winning a cup fixture in round 4 (round of 16) on league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = win cup fixture)							
Opponent Position	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.19***	0.19***	0.19***	0.19***	0.19***	0.18***	0.18***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Panel B: IV estimates (dependent variable = next league performance)							
Win Cup Fixture	1.03**	1.02**	1.00**	0.93*	0.94*	0.87	0.87
	(0.42)	(0.43)	(0.43)	(0.43)	(0.43)	(0.45)	(0.45)
F-Statistic	24.43	22.71	16.98	13.73	11.40	9.83	9.28
R^2	0.13	0.17	0.17	0.17	0.17	0.22	0.23
Observations	330	330	330	330	330	330	330

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 4. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Table 4.5, which investigates the impact of winning a cup fixture in Round 5 (Round of 32) on subsequent league performance, the first-stage results in Panel A confirm the robustness of the instrumental variables. Both opponent position and division are highly statistically significant at the 1% level, and the F-statistics remain strong across all models, ranging from 16.30 to 48.02, well above the threshold for weak instruments. This consistency across models indicates that the instruments are reliable predictors of the likelihood of winning the cup fixture.

The second-stage results in Panel B provide evidence that winning a cup fixture in Round 5 positively affects league performance, with the win coefficients ranging from 0.53 to 0.55. These effects are statistically significant at the 5% level in most models, showing that the treatment (winning the cup fixture) leads to an improvement in subsequent league performance. Compared to Round 4, the magnitude of the effect is smaller, suggesting that while winning continues to provide a boost, the psychological or momentum effects may vary depending on the stage of the competition.

Adding control variables like team position, travel distance, and recovery time does not significantly change the effect size, as the coefficients remain consistent across models. This robustness suggests that the positive effect of winning on league performance in Round 5 is not heavily affected by these added team characteristics. The results stay consistent for this round. However, when a broader set of controls is added in Models 6 and 7, the effect remains stable. These models incorporate team-specific attributes, including squad size and financial factors. Even after accounting for these deeper team qualities, the effect stays significant. This supports the idea that a positive causal link between winning the cup fixture and league performance exists.

The results for Round 5 align with those for Round 4, showing that the positive effect of winning carries over to the early rounds of the competition. The consistent significant coefficients across models indicate that winning in the Round of 32 reliably leads to better league performance, regardless of team strength, match specifics, or broader variables. This suggests that while factors like psychological momentum or morale might have some influence, the real benefits of winning, such as strategic gains or sharper focus, play an important role in league performance.

Table 4.5: Estimates of the effect of winning a cup fixture in round 5 (round of 32) on league performance using random drawn opponent position and division as an IV

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Panel A: First stage (dependent variable = win cup fixture)							
Opponent Position	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Opponent Division	0.17***	0.17***	0.17***	0.17***	0.17***	0.17***	0.18***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Panel B: IV estimates (dependent variable = next league performance)							
Win Cup Fixture	0.53*	0.53*	0.53*	0.55**	0.55**	0.54**	0.54**
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)
F-Statistic	48.02	38.00	28.52	22.90	19.07	16.45	16.30
R^2	0.12	0.14	0.14	0.14	0.14	0.18	0.19
Observations	705	705	705	705	705	705	705

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 5. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Table 4.6, we analyze the effect of winning a cup fixture in Round 6 (Round of 64) on subsequent league performance. The first-stage results in Panel A continue to show strong and statistically significant relationships between the opponent position and division and the likelihood of winning the cup fixture, with F-statistics well above the threshold for weak instruments. These F-statistics range from 14.24 to 39.91, ensuring confidence in the relevance of the instruments used to predict winning.

However, in contrast to the earlier rounds (Round 4 and Round 5), the second-stage results in Panel B show no clear statistically significant effect of winning a cup fixture on league performance. While the coefficient on winning a cup fixture is positive across all models, it fails to achieve statistical significance in most specifications, with the exception of Model 5 where the effect is significant at the 10% level. Even in this model, the coefficient is small (0.15) and does not represent a strong or highly robust causal effect. As more controls are added in Models 6 and 7, the coefficient decreases further and loses all statistical significance. This suggests that any potential effect of winning the cup fixture diminishes once more detailed team characteristics are accounted for.

These findings imply that, unlike in other earlier rounds where winning a cup fixture was more closely associated with subsequent league performance, in the Round of 64, the effect is smaller. The results suggest that for larger, stronger teams competing in earlier rounds of the cup, winning does not necessarily translate into a boost in league performance. This could be due to the fact that stronger teams might be expected to win in these early stages, meaning that the boost from winning is already priced into their expected league performance, providing the additional impact of the win insignificant.

The implications of these findings are important: the lack of a strong causal effect in Round 6 could reflect the lower importance of early-stage cup wins, particularly for well-resourced teams. These results also show that winning in the crucial rounds, specifically in Rounds 4 and 5, which are the Round of 16 and Round of 32, can positively impact league outcomes. However, earlier wins may not have the same impact once team characteristics are taken into account.

For Rounds 1, 2, and 3, the sample size might be insufficient to detect causal effects, indicating that further research with more data is needed to draw reliable conclusions.

Table 4.6: Estimates of the effect of winning a cup fixture in round 6 (round of 64) on league performance using random drawn opponent position and division as an IV

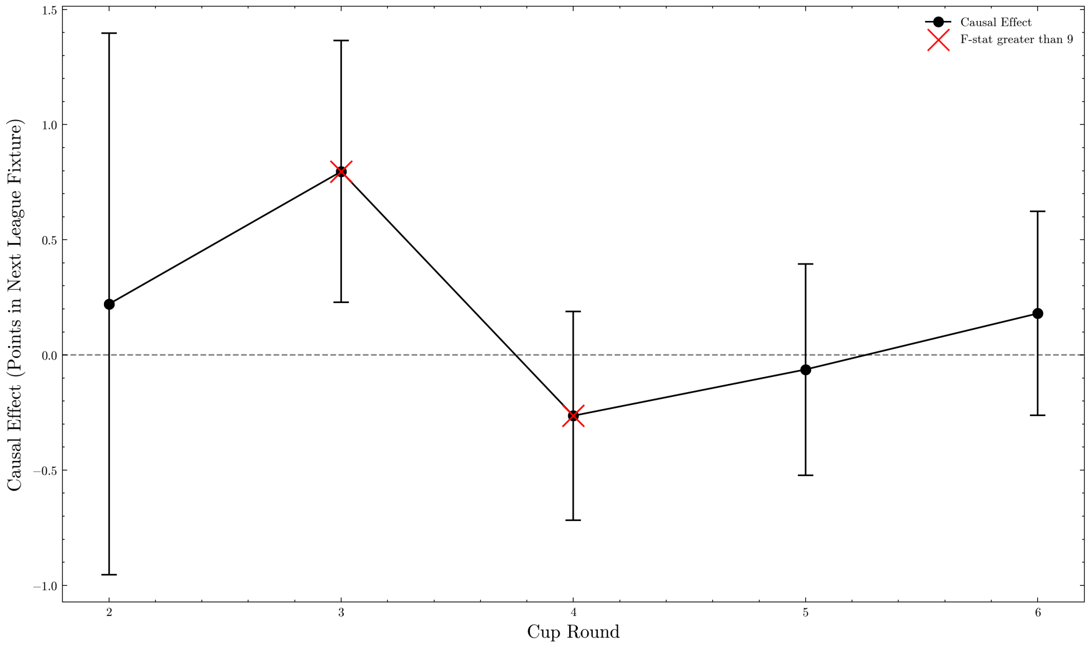
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Panel A: First stage (dependent variable = win cup fixture)							
Opponent Position	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.02***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Opponent Division	0.16***	0.16***	0.16***	0.16***	0.16***	0.16***	0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Panel B: IV estimates (dependent variable = next league performance)							
Win Cup Fixture	0.13	0.14	0.14	0.14	0.15*	0.12	0.09
	(0.31)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.31)
F-Statistic	39.91	32.84	24.70	19.77	16.59	14.24	15.14
R^2	0.11	0.14	0.14	0.14	0.14	0.17	0.20
Observations	618	618	618	618	618	618	618

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 6. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 The effect of cup participation

In this section the effect of participating in a cup round on the effect of points obtained in the league game within the first 5 days of participating in the cup round is examined. Figure 4.2 shows a summary of the results of the causal effects analysis. In none of the cup rounds there are significant causal effects for cup participation on domestic league performance. However, in Rounds 3 and 4, non-significant causal effects with stronger instrument relevance have been found.

Figure 4.2: Causal Effect of Participation in a Domestic Cup Round on League Performance.



Note: This figure illustrates a summary of the estimated causal effects of participating in a domestic cup round on league performance, measured in points obtained in the subsequent league fixture. Each point represents the estimated effect for a specific cup round (1 through 6), shown with confidence intervals. The black line indicates the main effect, while the red crosses highlight rounds where the first-stage F-statistic exceeds 9, suggesting stronger instrument relevance.

The results for the effect of cup participation on subsequent league performance across all rounds show no significant causal effects. In Round 5 (Table 5.2), the IV estimates in Panel B show cup participation coefficients ranging from -0.06 to 0.05, none of which are statistically significant. The standard errors remain large, suggesting high variability in the estimates, and the F-statistics for the first-stage regression in Panel A indicate that the instruments are strong, with values ranging from 6.06 to 20.27. This suggests that the instrument performs well for some model specifications in predicting cup participation, but participation itself does not have a statistically significant effect on performance in the next league fixture.

Similarly, for other rounds (as seen in the Appendix tables for Rounds 2, 3, 4, and 6), there is no statistically significant effect of cup participation on league performance. In Round 2 (Table E.1), the cup participation coefficients in Panel B range from 0.01 to 0.36, but none of these results are significant. The F-statistics for the first-stage regression range from 3.73 to 7.73, suggesting a weak instrument and no significant outcome effect. In Round 3 (Table E.2),

Table 4.7: Estimates of the effect of cup participation in round 5 on subsequent league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Opponent Position	0.01*	0.01*	0.01*	0.01**	0.01**	0.01**	0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.18***	0.18***	0.18***	0.17***	0.17***	0.19***	0.18***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Participation	-0.06	-0.06	0.01	0.05	-0.07	-0.05	-0.06
	(0.47)	(0.47)	(0.46)	(0.49)	(0.49)	(0.45)	(0.46)
F-Statistic	20.27	14.06	10.64	8.87	7.38	6.56	6.06
R^2	0.15	0.15	0.15	0.16	0.16	0.20	0.21
Observations	240	240	240	240	240	240	240

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 5. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

despite a relatively stronger first-stage (F-statistics ranging from 9.48 to 20.00), the second-stage results show no significant impact of cup participation on league performance, with coefficients ranging from 0.70 to 0.81.

The results from Round 4 (Table E.3) continue to show no significant effects, with the IV estimates in Panel B ranging from -0.08 to -0.26 and large standard errors, indicating high uncertainty in the estimates. The F-statistics in the first stage range from 9.22 to 24.30, indicating that the instrument is stronger in this round. Round 6 (Table E.4) follows the same pattern, with coefficients ranging from 0.08 to 0.19 in Panel B, none of which are statistically significant. The first-stage F-statistics range from 7.81 to 19.47, again indicating weak to moderately strong instruments.

In summary, across all rounds, while instrumental variables are often reasonably strong predictors of cup participation, participation itself does not have a significant effect on subsequent league performance. This is in contrast to the significant effects observed when analyzing cup wins, particularly in earlier stages like Round 4 and Round 5, where cup wins were associated with a significant positive effect on league performance.

Chapter 5

Robustness

5.1 Distance travelled as IV

In this section, we compare the performance of two instruments: opponent position and division with distance traveled as IV for analyzing the effect of winning a cup round on subsequent league performance. The first-stage results using opponent position and division show strong and consistent F-statistics, particularly in Round 4 and Round 5, with values ranging from 9.28 to 24.43 (see Table 4.4), confirming that these instruments are effective in predicting cup participation and wins. The coefficients for opponent division are generally positive and significant, indicating that teams facing lower-division opponents are more likely to win, which aligns with the expected difference in strength.

On the other hand, using distance traveled as an instrument presents weaker F-statistics in the early rounds, with values such as 1.54 in Round 4 but improving significantly to over 8.39 in later models and rounds (Table E.3). The first-stage coefficients for distance traveled are negative, reflecting the intuitive impact of fatigue, suggesting that longer travel distances decrease the likelihood of winning a cup fixture. However, the second-stage results for distance traveled as an IV are generally insignificant, indicating that while travel distance may capture fatigue, it does not robustly explain the causal effect of cup wins on league performance. In contrast, opponent position and division provide clearer, statistically significant second-stage results in the later rounds, particularly in Round 5 (Table 4.5), making them more reliable instruments for this context.

Table 5.1: Estimates of the effect of cup participation in round 4 on subsequent league performance using Distance Traveled as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Distance Traveled	-0.00024	-0.00028*	-0.00035*	-0.00034*	-0.00034*	-0.00035**	-0.00034*
	(0.00020)	(0.00019)	(0.00018)	(0.00018)	(0.00018)	(0.00018)	(0.00018)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Win	0.60	0.70	0.70	0.58	0.63	0.40	0.42
	(2.23)	(1.93)	(1.55)	(1.58)	(1.58)	(1.53)	(1.57)
F-Statistic	1.54	8.39**	16.72***	12.89***	10.28***	9.26***	8.63***
R^2	0.00	0.01	0.02	0.03	0.03	0.07	0.07
Observations	332	332	332	332	332	332	332

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 4. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5.2: Estimates of the effect of cup participation in round 5 on subsequent league performance using Distance Traveled as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Distance Traveled	-0.00018**	-0.00018**	-0.00022***	-0.00022***	-0.00022***	-0.00020***	-0.00020***
	(0.00009)	(0.00008)	(0.00008)	(0.00008)	(0.00008)	(0.00008)	(0.00008)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Win	-0.21	-0.22	0.05	-0.05	-0.07	-0.21	-0.22
	(1.22)	(1.24)	(0.99)	(0.99)	(0.99)	(1.09)	(1.10)
F-Statistic	4.45**	8.24***	21.15***	15.83***	12.85***	9.45***	9.11***
R^2	0.01	0.04	0.13	0.13	0.14	0.16	0.17
Observations	414	414	414	414	414	414	414

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 5. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The comparison highlights that opponent position and division better capture competitive dynamics, while distance traveled may reflect logistical challenges but might not have the predictive power for league performance outcomes.

5.2 Heterogeneity analysis

As Dutch football legend Johan Cruijff famously stated, "Ik heb een zak geld nog nooit een doelpunt zien scoren" ("I've never seen a bag of money score a goal"), emphasizing that success on the pitch is not only about financial strength. This sentiment aligns with our findings on the causal effects of domestic cup participation, where we examine the impact of differences in total market values between high- and low-value teams.

Table 5.3 presents the 2SLS IV estimates for teams in the top 20% and bottom 20% of market value. The results for teams with high market value demonstrate minimal and largely insignificant effects in the early rounds of the cup. For example, winning a fixture in Round 4 yields a insignificant effect of 1.52 on subsequent league performance for the top 20% market value teams. However, in the lower market value teams, winning a cup fixture in the same round has a stonger effect, with statistically significant estimates, such as a 0.99 increase in points in the next league fixture. These findings suggest that, contrary to expectations, teams with lower market value experience stronger league performance following cup wins, particularly in the earlier rounds. This could indicate that teams with fewer resources are more likely to leverage the momentum of a cup win in the next league fixture.

The F-statistics further strengthens the robustness of the model for lower market value teams, consistently exceeding the critical value of 10 in most rounds. This contrasts with the weaker identification seen among the teams with higher total market values, particularly in earlier rounds, where F-statistics are around 2.93 for Round 5 and drop further for Round 6.

Another discussed topic in football management is the relationship between squad size and team performance in football. Larger squads allow for more rotation opportunities, particularly in congested fixture schedules, while smaller squads may struggle to cope with the physical and tactical demands of competing in multiple competitions.

Table 5.4 presents the IV estimates for the top 20% and bottom 20% of teams based on squad size. For teams with large squads, winning a cup fixture in Round 4 has a strong and statistically significant effect on subsequent league performance, with an IV estimate of 3.10

Table 5.3: Estimates of the effect of winning a cup fixture on subsequent league performance using Opponent Position and Division as IV for high and low total market values of teams

	Total Market Value					
	Top 20%			Bottom 20%		
Cup Round	4	5	6	4	5	6
Panel A: First Stage						
Division	0.05	0.17**	0.06	0.19**	0.26***	0.15***
Opponent Position	0.01	0.00	0.03***	0.03**	0.02*	0.02**
Panel B: IV estimates						
Cup Win	1.52	-0.29	-0.41	0.99**	2.39**	1.49**
F-Stat	1.46	2.93***	2.64***	1.08	4.01***	3.10***
Observations	82	93	69	50	49	98
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Model specification	7	7	7	7	7	7

Note: The table reports both the first-stage and second-stage 2SLS IV estimates with full control variables for the top 20% and bottom 20% market value teams in rounds 4, 5, and 6 of domestic cup competitions. The control variables include team position, cup fixture related variables, recovery time, and team characteristics.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

($p < 0.01$). This suggests that, as expected, larger squads can handle the demands of playing in mulitple competitions better. This is likely due to their ability to rotate players and lower fatigue of the squad. In Rounds 5 and 6, the effect size decreases slightly, but remain positive and significant in Round 6 with an estimate of 1.40 ($p < 0.10$).

In contrast, for teams with smaller squads, the results are more mixed. Winning a cup fixture in Round 4 shows a moderate effect, with an IV estimate of 1.33, but without statistical significance. By Rounds 5 and 6, the effect turns negative, though not statistically significant. This could indicate that smaller squads face increasing difficulty in maintaining league performance as the cup competition progresses, possibly due to the physical impact.

The F-statistics confirm the robustness of these results for both large and small squads, consistently surpassing the threshold for strong instruments in most cases. The results suggest that larger squads are better positioned to maintain strong league performances following a cup fixture, while smaller squads may struggle.

Table 5.4: Estimates of the effect of winning a cup fixture on subsequent league performance using Opponent Position and Division as IV for large and small squad sizes of teams

	Team Size					
	Top 20%			Bottom 20%		
Cup Round	4	5	6	4	5	6
Panel A: First Stage						
Division	0.16**	0.09*	0.16***	0.22***	0.23***	0.12***
Opponent Position	-0.01	0.01	0.02**	0.02**	0.02**	0.03***
Panel B: IV estimates						
Cup Win	3.10***	1.82	1.40*	1.33	-0.02	-0.22
F-Stat	1.82	1.98**	3.39***	3.74***	4.52***	2.98***
Observations	79	83	92	76	94	129
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Model specification	7	7	7	7	7	7

Note: The table reports both the first-stage and second-stage 2SLS IV estimates with full control variables for the top 20% and bottom 20% squad size teams in rounds 4, 5, and 6 of domestic cup competitions. The control variables include team position, cup fixture related variables, recovery time, and team characteristics.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 6

Conclusion

This research investigates the causal effect of domestic cup participation and winning on subsequent league performance in European football, focusing on four major domestic cup competitions: the FA Cup (England), DFB Pokal (Germany), KNVB Beker (Netherlands) and Taça de Portugal (Portugal). By using the random assignment of cup draws as a plausibly exogenous factor, we estimate the impact of participating in and winning cup fixtures on league performance. Our results show clear and consistent patterns: winning a cup fixture significantly improves league performance in the next fixture, particularly for smaller teams. For larger teams, this effect is less pronounced. The most significant results were observed in the earlier rounds of the cup competition.

The fact that smaller teams benefit more from winning cup fixtures than larger teams raises some interesting questions. There are several possible reasons for this. Smaller teams might experience a psychological boost or use the momentum from a cup victory to perform better in their next league game. Larger teams, however, tend to have more resources and deeper squads, which might help them handle fixture congestion better and reduce the importance of any temporary morale boost from winning. This finding is similar to the research by Moffat (2020), which suggests that smaller teams benefit more from participating in European tournaments due to the exposure and resources gained. Although their definition of smaller teams is different, as it applies to European-level tournaments, our results support the idea that resource-limited teams feel the impact of wins more strongly. The findings of Cabras et al. (2022) also align with our results, showing that larger teams are less affected by the demands of playing in multiple competitions, especially those that include international fixtures.

However, simply participating in a cup round without winning does not appear to have a significant effect on subsequent league performance. This holds true for all rounds and teams, indicating that fixture congestion alone, without the psychological momentum of a win, does not affect performance. These findings align with previous research by Poli et al. (2015), which showed limited effects of participation in the Champions League on league performance, particularly for top-tier teams. Our study extends these insights to the context of the domestic cup.

Despite the robustness of our empirical strategy, there are limitations to our analysis. By focusing only on league fixtures played within five days of a cup fixture, we may overlook the longer-term effects of cup participation and winning. Additionally, our dataset includes only four

European countries, and the findings may not be generalized to leagues with different scheduling practices or cup formats, particularly in countries where cup draws are seeded, such as France, Italy, and Spain. Expanding this analysis to include more countries and different cup structures would help provide a broader understanding of the impact of domestic cup competitions on league performance.

Another limitation is the lack of detailed data on squad rotation, injuries, and key players in starting lineups. These factors may significantly influence how teams cope with fixture congestion, but are not fully captured in our analysis. Research by Carling and Dupont (2012) and Dupont and Wisløff (2010) have shown that fixture congestion can increase injury risks and reduce recovery time, potentially affecting team performance in ways we have not fully explained.

A major limitation of our study is the relatively short timeframe of our data, which only covers football seasons from 2012 onward for the FA Cup and DFB Pokal. This short period reduces the number of observations, especially for the later stages of the competition like the semi-finals and finals, where we could not perform a robust Two-Stage Least Squares (2SLS) analysis. This lack of data prevented us from finding significant causal effects in these key rounds. Access to a more extensive historical dataset, potentially covering 25 to 50 years, would provide a more complete view of cup participation's impact, especially in the later rounds. Although data on the FA Cup goes back to 1871, it is not freely available from our provider. Expanding the dataset would improve the robustness of future analyses and offer a clearer understanding of fixture congestion's long-term effects across different eras.

To further demonstrate the robustness of our findings and validate the exogeneity assumption, incorporating a placebo test would be beneficial. A practical approach for this research could involve assessing the impact of winning or participating in a cup fixture on a subsequent league fixture occurring beyond the immediate five-day window. For example, we could examine fixtures that take place two to three weeks after a cup fixture, where we would not expect any direct impact. If our identification strategy holds, these placebo tests should show no significant results. This outcome would confirm that the observed positive effects on league performance are specific to the immediate aftermath of a cup fixture and not influenced by unobserved confounding factors or selection bias. Including such a test would strengthen our argument that the significant effects found, especially for smaller teams, are indeed causal.

Future research could extend our analysis by examining metrics such as distance covered and injury rates in the next league fixture to assess fixture congestion from a broader perspective. This would provide deeper insights into the physical impact on players and help inform UEFA, FIFA, and football teams on managing congested schedules more effectively. Such findings could lead to better scheduling practices, reduced player fatigue, and improved competition sustainability.

Another extension study could involve examining the cumulative effects of participation across all rounds of domestic cup competitions. The factorial IV methodology proposed by Blackwell and Pashley (2021) presents a promising approach, to apply it to our study we would treat each cup round as a separate experimental unit to estimate the cumulative impact on league performance. However, in our current setup, this methodology would not be effective due to the limited sample size, which is insufficient to produce reliable results for factorial experiments. A more extensive dataset including more seasons would be needed to ensure robust analysis.

Taken together, our findings add to the existing research on fixture congestion by showing that winning a domestic cup match has a positive and significant effect on subsequent league performance, particularly for smaller clubs. Larger teams, on the other hand, appear more capable of handling the physical and tactical challenges associated with participating in multiple competitions at the same time. These results have practical implications for football managers, especially those leading smaller clubs, as they highlight the importance of leveraging momentum gained from cup victories during congested periods of the season. Our study emphasizes the strategic benefits of domestic cup successes and offers valuable insights for decision-making in football management.

Appendix A

Variable Definitions

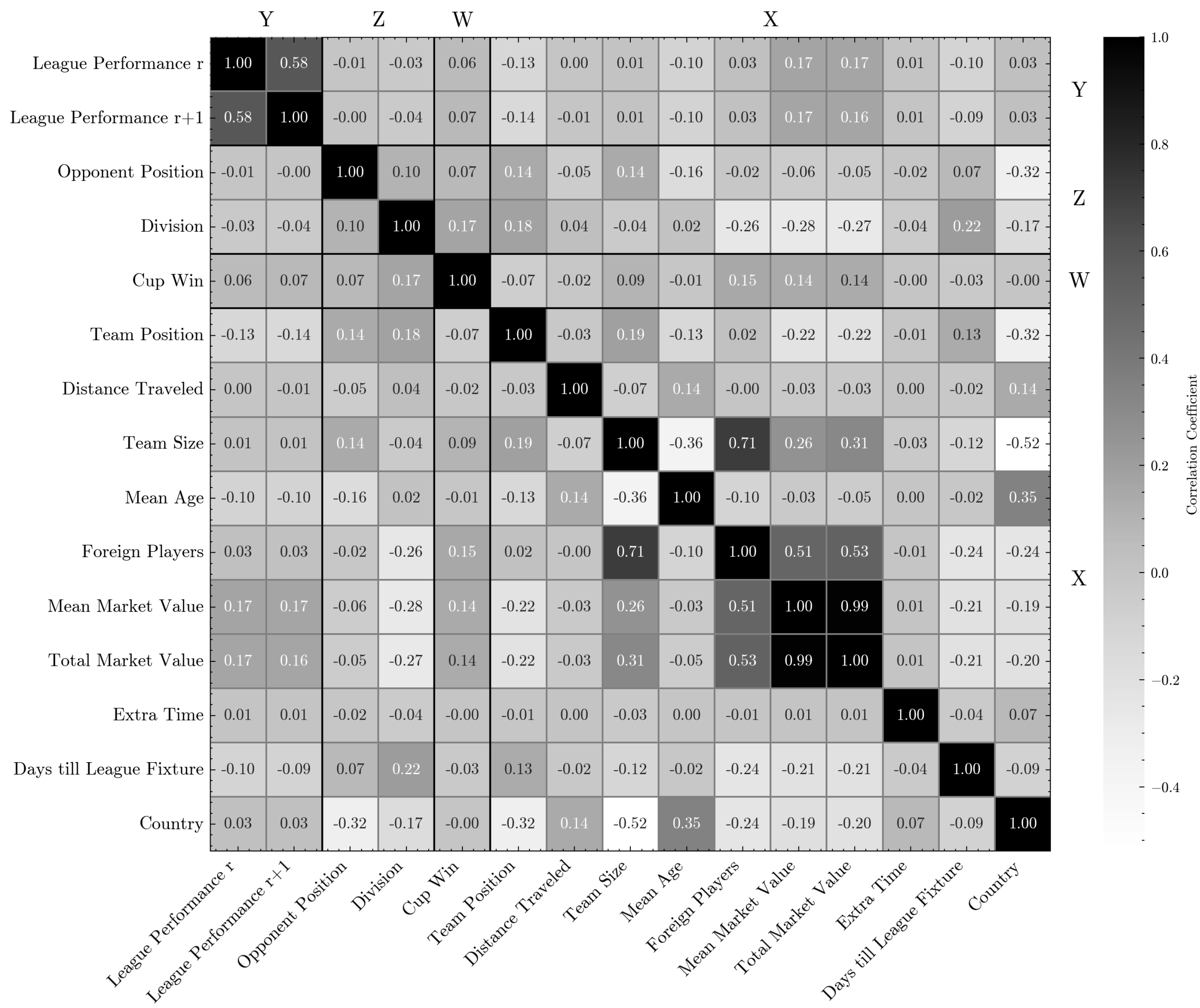
Table A.1: Variable Definitions

Variable	Type	Definition
League Performance	Outcome	Points earned by the team in the league match following a cup match, with points assigned as: 3 for a win, 1 for a draw, and 0 for a loss.
Opponent Position	Instrument	Opponent's league rank in the previous season, serving as an exogenous factor influencing the likelihood of winning the cup match.
Division	Instrument	The league division in which the opponent competed in the prior season, providing an additional measure of opponent strength.
Cup Win	Treatment	Binary indicator for whether the team won the cup match (1 = win, 0 = loss or draw), used to assess the impact on subsequent league performance.
Team Position	Control	Team's rank in their domestic league in the previous season, used to control for baseline team strength.
Team National Position	Control	Unified ranking of the team across all divisions, providing a consistent measure of team strength across divisions.
Team Size	Control	Total number of players in the team, which may relate to squad depth and capacity to handle fatigue.
Foreign Players	Control	Count of foreign players in the team, potentially affecting style of play and performance dynamics.
Mean Age	Control	Average age of the team, as a measure of player experience and squad maturity.
Mean Market Value	Control	Average market value of the team's players, serving as a proxy for skill level and overall quality.
Total Market Value	Control	Combined market value of all players, representing team financial capacity and talent level.
Distance Traveled	Control	Distance covered by the team to the cup match venue, with greater distances possibly contributing to fatigue and affecting performance.
Extra Time	Control	Binary indicator of whether the cup match extended into extra time, which may add physical strain affecting future games.
Days till League Fixture	Control	Number of days between the cup and subsequent league matches, controlling for the available recovery period.
Country Code	Control	Numeric code representing the team's country, adjusting for country-specific league structure and competitiveness. Lower codes indicate stronger leagues as per UEFA rankings.

Appendix B

Correlations

Figure B.1: Correlation heatmap of league performance, cup performance, and team characteristics.



Note: This heatmap visualizes the correlations among key variables in our analysis. Variables are categorized as follows: Y (Outcome Variables) represent league performance; Z (Instrumental Variables) represent opponent strength and division; W (Treatment Variable) indicates a win in the cup fixture; and X (Control Variables) include factors influencing outcomes. The observed strong correlations suggest that financial metrics and squad characteristics impact performance.

Appendix C

Model specifications

A.1 Win Cup Fixture

Model 1:

First Stage: Baseline model, estimating the probability of winning a cup fixture based on instrumental variables only.

$$\text{Cup Win}_i = \alpha_1 + \gamma_1 \cdot \text{Opponent Position}_i + \gamma_2 \cdot \text{Division}_i + \epsilon_i \tag{A.1}$$

Second Stage: Estimating the effect of winning the cup fixture on league performance.

$$\text{League Performance}_i = \alpha_2 + \beta \cdot \widehat{\text{Cup Win}}_i + \nu_i \tag{A.2}$$

Model 2:

First Stage: Predicting cup win likelihood with added team position.

$$\begin{aligned} \text{Cup Win}_i = \alpha_1 + \gamma_1 \cdot \text{Opponent Position Previous}_i + \gamma_2 \cdot \text{Division}_i \\ + \delta_1 \cdot \text{Team Position}_i + \epsilon_i \end{aligned} \tag{A.3}$$

Second Stage: Assessing the impact of cup win on league performance, adjusting for team position.

$$\text{League Performance}_i = \alpha_2 + \beta \cdot \widehat{\text{Cup Win}}_i + \lambda_1 \cdot \text{Team Position}_i + \nu_i \tag{A.4}$$

Model 3:

First Stage: Adding distance traveled as a factor influencing cup win probability.

$$\begin{aligned} \text{Cup Win}_i = \alpha_1 + \gamma_1 \cdot \text{Opponent Position Previous}_i + \gamma_2 \cdot \text{Division}_i \\ + \delta_1 \cdot \text{Team Position}_i + \delta_2 \cdot \text{Distance}_i + \epsilon_i \end{aligned} \tag{A.5}$$

Second Stage: Including distance traveled in the model for league performance.

$$\begin{aligned}\text{League Performance}_i = \alpha_2 + \beta \cdot \widehat{\text{Cup Win}}_i + \lambda_1 \cdot \text{Team Position}_i \\ + \lambda_2 \cdot \text{Distance}_i + \nu_i\end{aligned} \tag{A.6}$$

Model 4:

First Stage: Adding recovery time as a factor.

$$\begin{aligned}\text{Cup Win}_i = \alpha_1 + \gamma_1 \cdot \text{Opponent Position}_i + \gamma_2 \cdot \text{Division}_i \\ + \delta_1 \cdot \text{Team Position}_i + \delta_2 \cdot \text{Distance}_i \\ + \delta_3 \cdot \text{Days till League Fixture}_i + \epsilon_i\end{aligned} \tag{A.7}$$

Second Stage: Including recovery time in the model for league performance.

$$\begin{aligned}\text{League Performance}_i = \alpha_2 + \beta \cdot \widehat{\text{Cup Win}}_i + \lambda_1 \cdot \text{Team Position}_i \\ + \lambda_2 \cdot \text{Distance}_i + \lambda_3 \cdot \text{Days till League Fixture}_i + \nu_i\end{aligned} \tag{A.8}$$

Model 5:

First Stage: Adding extra time as an additional factor.

$$\begin{aligned}\text{Cup Win}_i = \alpha_1 + \gamma_1 \cdot \text{Opponent Position}_i + \gamma_2 \cdot \text{Division}_i \\ + \delta_1 \cdot \text{Team Position}_i + \delta_2 \cdot \text{Distance}_i \\ + \delta_3 \cdot \text{Days till League Fixture}_i + \delta_4 \cdot \text{Extra Time}_i + \epsilon_i\end{aligned} \tag{A.9}$$

Second Stage: Assessing the impact of extra time on league performance.

$$\begin{aligned}\text{League Performance}_i = \alpha_2 + \beta \cdot \widehat{\text{Cup Win}}_i + \lambda_1 \cdot \text{Team Position}_i \\ + \lambda_2 \cdot \text{Distance}_i + \lambda_3 \cdot \text{Days till League Fixture}_i \\ + \lambda_4 \cdot \text{Extra Time}_i + \nu_i\end{aligned} \tag{A.10}$$

Model 6:

First Stage: Adding team-specific characteristics.

$$\begin{aligned}\text{Cup Win}_i = \alpha_1 + \gamma_1 \cdot \text{Opponent Position}_i + \gamma_2 \cdot \text{Division}_i \\ + \delta_1 \cdot \text{Team Position}_i + \delta_2 \cdot \text{Distance}_i \\ + \delta_3 \cdot \text{Days till League Fixture}_i + \delta_4 \cdot \text{Extra Time}_i \\ + \delta_5 \cdot \text{Mean Age}_i + \delta_6 \cdot \text{Team Size}_i \\ + \delta_7 \cdot \text{Total Market Value}_i + \epsilon_i\end{aligned} \tag{A.11}$$

Second Stage: Assessing the effect of team-specific characteristics on league performance.

$$\begin{aligned}\text{League Performance}_i = \alpha_2 &+ \beta \cdot \widehat{\text{Cup Win}}_i + \lambda_1 \cdot \text{Team Position}_i \\ &+ \lambda_2 \cdot \text{Distance}_i + \lambda_3 \cdot \text{Days till League Fixture}_i \\ &+ \lambda_4 \cdot \text{Extra Time}_i + \lambda_5 \cdot \text{Team Size}_i \\ &+ \lambda_6 \cdot \text{Mean Age}_i + \lambda_7 \cdot \text{Total Market Value}_i \\ &+ \lambda_8 \cdot \text{Foreign Players}_i + \nu_i \end{aligned} \tag{A.12}$$

Model 7:

First Stage: Adding team-specific characteristics.

$$\begin{aligned}\text{Cup Win}_i = \alpha_1 &+ \gamma_1 \cdot \text{Opponent Position}_i + \gamma_2 \cdot \text{Division}_i \\ &+ \delta_1 \cdot \text{Team Position}_i + \delta_2 \cdot \text{Distance}_i \\ &+ \delta_3 \cdot \text{Days till League Fixture}_i + \delta_4 \cdot \text{Extra Time}_i \\ &+ \delta_5 \cdot \text{Mean Age}_i + \delta_6 \cdot \text{Team Size}_i \\ &+ \delta_7 \cdot \text{Total Market Value}_i + \epsilon_i \end{aligned} \tag{A.13}$$

Second Stage: Assessing the effect of team-specific characteristics on league performance.

$$\begin{aligned}\text{League Performance}_i = \alpha_2 &+ \beta \cdot \widehat{\text{Cup Win}}_i + \lambda_1 \cdot \text{Team Position}_i \\ &+ \lambda_2 \cdot \text{Distance}_i + \lambda_3 \cdot \text{Days till League Fixture}_i \\ &+ \lambda_4 \cdot \text{Extra Time}_i + \lambda_5 \cdot \text{Mean Age}_i \\ &+ \lambda_6 \cdot \text{Team Size}_i + \lambda_7 \cdot \text{Total Market Value}_i \\ &+ \delta_8 \cdot \text{Country Code}_i + \nu_i \end{aligned} \tag{A.14}$$

A.2 Participation Cup Fixture

Follow a similar structure for the Participation Cup Fixture models, replacing the Cup Win treatment variable with Cup Participation, as defined in Chapter 2.

Appendix D

Tests for instrument independence

This appendix presents the Chi-Squared test results for independence by round for the domestic cup competitions in Portugal, the Netherlands, and Germany. These results provide further support for the randomness assumption applied in the empirical strategy.

Table D.1: Chi-Squared Test for Independence in the German DFB Pokal

Round	**Fixture vs Division**			**Fixture vs Position**		
	Chi-2	**p-value**	**df**	**Chi-2**	**p-value**	**df**
1	11.48	.4041	11	108.00	.5361	110
2	27.69	.2276	23	324.00	.4582	322
3	90.06	.5958	94	795.17	.5316	799
4	191.68	.4522	190	1673.31	.7325	1710

Table D.2: Chi-Squared Test for Independence in the Dutch KNVB Beker

Round	**Fixture vs Division**			**Fixture vs Position**		
	Chi-2	**p-value**	**df**	**Chi-2**	**p-value**	**df**
1	0.00	1.0000	0	42.00	.4710	42
2	14.48	.4893	15	191.25	.2690	180
3	30.56	.4375	30	493.08	.3300	480
4	58.84	.5547	61	1063.05	.2803	1037
5	99.47	.6339	105	1928.47	.8542	1995
6	164.85	.4229	162	3060.08	.5872	3078

Table D.3: Chi-Squared Test for Independence by in the Taça de Portugal

Round	**Fixture vs Division**			**Fixture vs Position**		
	Chi-2	**p-value**	**df**	**Chi-2**	**p-value**	**df**
1	7.47	.3820	7	32.00	.6137	35
2	23.83	.4131	23	201.14	.1836	184
3	33.56	.2988	30	437.97	.6614	450
4	100.50	.5789	104	1956.17	.6202	1976
5	165.03	.7652	179	3543.11	.6662	3580
6	94.70	.5183	96	1815.09	.5544	1824

Appendix E

IV estimates cup participation

Table E.1: Estimates of the effect of cup participation in round 2 (semifinal) on subsequent league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Opponent Rank	0.03*	0.03**	0.03**	0.03**	0.03*	0.03**	0.04***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.39	0.43	0.40	0.43	0.44	0.25	0.18
	(0.29)	(0.25)	(0.25)	(0.24)	(0.25)	(0.23)	(0.23)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Participation	0.01	0.11	0.29	0.33	0.36	-0.08	0.22
	(1.11)	(0.99)	(1.02)	(1.01)	(1.03)	(1.33)	(1.18)
F-Statistic	3.73	7.73	6.06	5.38	4.34	5.23	5.13
R^2	0.20	0.45	0.47	0.51	0.51	0.68	0.71
Observations	32	32	32	32	32	32	32

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 2 (semifinal). Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.2: Estimates of the effect of cup participation in round 3 on subsequent league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Opponent Rank	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.14**	0.12**	0.12**	0.11*	0.12**	0.16***	0.16***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Participation	0.86	0.70	0.77	0.81	0.72	0.81	0.80
	(0.55)	(0.57)	(0.57)	(0.61)	(0.63)	(0.57)	(0.57)
F-Statistic	14.82	20.00	14.95	12.92	11.17	10.26	9.48
R^2	0.15	0.26	0.26	0.28	0.29	0.36	0.37
Observations	172	172	172	172	172	172	172

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 3. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.3: Estimates of the effect of cup participation in round 4 on subsequent league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Opponent Rank	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Opponent Division	0.22***	0.22***	0.22***	0.22***	0.22***	0.22***	0.21***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Participation	-0.08	-0.16	-0.16	-0.17	-0.17	-0.09	-0.26
	(0.42)	(0.42)	(0.42)	(0.43)	(0.43)	(0.42)	(0.45)
F-Statistic	24.30	20.06	15.05	13.79	11.46	9.39	9.22
R^2	0.16	0.19	0.19	0.21	0.21	0.25	0.27
Observations	266	266	266	266	266	266	266

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 4. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.4: Estimates of the effect of cup participation in round 6 on subsequent league performance using random drawn opponent position and division as an IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: First stage (dependent variable = cup participation)							
Opponent Rank	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Opponent Division	0.14***	0.14***	0.14***	0.14***	0.14***	0.14***	0.14***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Panel B: IV estimates (dependent variable = next league performance)							
Cup Participation	0.19	0.08	0.10	0.10	0.10	0.19	0.18
	(0.44)	(0.45)	(0.45)	(0.46)	(0.46)	(0.45)	(0.44)
F-Statistic	19.47	17.61	13.22	13.88	11.60	7.81	8.34
R^2	0.11	0.15	0.15	0.19	0.19	0.19	0.22
Observations	310	310	310	310	310	310	310

Note: The table reports both the first-stage and second-stage 2SLS IV estimates for Round 6. Standard errors are in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix F

Programming Code

All code used in this thesis was written in Python and is accessible on my GitHub page. The code contains every stage of the analysis, including:

- Data collection
- Exploratory data analysis
- Data preprocessing
- Causal effect analysis

For reproducibility and transparency, the complete codebase is available at:
`https://github.com/jellewillekes/causal-effect-cup-football`

References

Abadie, A. & Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*, *10*, 465–503.

Angrist, J. D., Imbens, G. W. & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

API-Football. (2024). *API-Football: Football statistics and data API.* Retrieved from `https://www.api-football.com/` (Accessed: 2024-10-31)

Blackwell, M. & Pashley, N. E. (2021). Noncompliance and instrumental variables for 2k factorial experiments. *Journal of the American Statistical Association*, *00*(0), 1–13. doi: 10.1080/01621459.2021.1978468

Cabras, S., Delogu, M. & Tena, J. d. D. (2022). Forced to play too many matches? a deep-learning assessment of crowded schedule. *Applied Economics*, *55*(52), 6187–6204.

Carling, L. G. F., C. & Dupont, G. (2012). Are physical performance and injury risk in a professional soccer team in match-play affected over a prolonged period of fixture congestion? *International Journal of Sports Medicine*, *33*(3), 36–42. doi: 10.1055/s-0031-1308946

Cea, S., Durán, G., Guajardo, M., Siebert, J. & Zamorano, G. (2020, Mar). An analytics approach to the fifa ranking procedure and the world cup final draw. *Annals of Operations Research*, *286*, 119–146. Retrieved from `https://doi.org/10.1007/s10479-019-03261-8` doi: 10.1007/s10479-019-03261-8

Colman, J. (2023). Rodri: Manchester city midfielder says players 'close' to striking over fixture congestion. *BBC Sport*. Retrieved from `https://www.bbc.com/sport/football/articles/cx20lg01mndo` (Accessed: 2024-09-18)

Dagaev, D. & Suzdaltsev, A. (2018, Jan). Competitive intensity and quality maximizing seedings in knock-out tournaments. *Journal of Combinatorial Optimization*, *35*, 170–188. Retrieved from `https://doi.org/10.1007/s10878-017-0164-7` doi: 10.1007/s10878-017-0164-7

Dellal, L.-P. C. R. E. C. K., A. & Orhant, E. (2015). The effects of a congested fixture period on physical performance, technical activity and injury rate during matches in a professional soccer team. *British Journal of Sports Medicine*, *49*(6), 390–394. doi: 10.1136/bjsports-2012-091290

Dupont, N.-M. M. A. M. D. B. S., G. & Wisløff, U. (2010). Effect of 2 soccer matches in a week on physical performance and injury rate. *The American Journal of Sports Medicine*, *38*(9), 1752–1758. doi: 10.1177/0363546510361236

FIFPro. (2020). The congested football calendar and its impact on player performance. *FIFPro*. Retrieved from `https://www.fifpro.org/en/rights/international-match`

-calendar/ (Accessed: 2024-04-11)

Fox Sports. (2022). Jose mourinho latest premier league coach to complain about fixture congestion. *Fox Sports*. Retrieved from https://amp.foxsports.com/stories/soccer/jose-mourinho-latest-premier-league-coach-to-complain-about-fixture-congestion (Accessed: 2024-02-29)

Franck, E. & Nüesch, S. (2012). Talent and/or popularity: What does it take to be a superstar? *Economic Inquiry*, *50*(1), 202–216.

Heckman, J. J. & Vytlacil, E. J. (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, *73*(3), 669–738.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945–960.

Imbens, G. W. (2014). Instrumental variables: an econometrician's perspective. *Statistical Science*, *29*(3), 323–358.

Imbens, G. W. & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Johnson, R. A. & Wichern, D. W. (2018). *Applied multivariate statistical analysis* (6th ed.). Upper Saddle River, NJ: Pearson.

Krumer, A. & Lechner, M. (2018). Midweek effect on soccer performance: Evidence from the german bundesliga. *Economic Inquiry*, *56*(1), 193–207.

L. Pillay, D. C. J. v. R. G. M. K. . V. G., D. Burgess. (2022). The congested international match calendar in football: views of 1055 professional male players. *BMC Sports Science, Medicine and Rehabilitation*, *14*(1), 200. Retrieved from https://doi.org/10.1186/s13102-022-00597-w doi: 10.1186/s13102-022-00597-w

Moffat, J. (2020). The impact of participation in pan-european competition on domestic performance in association football. *European Sport Management Quarterly*, *20*(4), 440–457.

Mogstad, M., Torgovitsky, A. & Walters, C. (2021, 11). The causal interpretation of two-stage least squares with multiple instrumental variables. *American Economic Review*, *111*, 3663-3698. doi: 10.1257/aer.20190221

Neyman, J. (1923). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, *5*, 463-480.

Nichols, M. W. (2014). The impact of visiting team travel on game outcome and biases in nfl betting markets. *Journal of Sports Economics*, *15*(1), 78–96.

Oberhofer, H., Philippovich, T. & Winner, H. (2010). Distance matters in away games: Evidence from the german football league. *Journal of Economic Psychology*, *31*(2), 200–211.

Olley, D. (2020). Manchester city boss pep guardiola says congested fixture schedule is 'too much' for players. *Sky Sports*. Retrieved from https://www.skysports.com/football/news/11679/12112506/manchester-city-boss-pep-guardiola-says-congested-fixture-schedule-is-too-much-for-players (Accessed: 2024-02-29)

Pearce, J. (2024). Liverpool's alisson criticises new champions league format over welfare concerns: 'players need to be listened to'. *The New York Times*. Retrieved from https://www.nytimes.com/athletic/5771839/2024/09/16/alisson-liverpool-champions-league/ (Accessed: 2024-09-18)

Poli, R., Besson, R. & Ravenel, L. (2015). Impact of uefa club competitions on domestic league results. *Football Observatory, International Centre for Sports Studies (CIES), University of Neuchatel*.

Ribeiro, M. T., Singh, S. & Guestrin, C. (2016). "why should I trust you?": Explaining the predictions of any classifier. *CoRR*, *abs/1602.04938*. Retrieved from http://arxiv.org/abs/1602.04938

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688.

Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, *75*(371), 591–593. Retrieved from https://doi.org/10.2307/2287653 doi: 10.2307/2287653

Scoppa, V. (2015). Fatigue and team performance in soccer: Evidence from the fifa world cup and the uefa european championship. *Journal of Sports Economics*, *16*(5), 482–507. doi: 10.1177/1527002513502794

Simmons, R. & Walker, I. (2010). Can subsidies help buy success? revenue sharing in english football. In *Optimal strategies in sports economics and management* (pp. 171–185). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from https://doi.org/10.1007/978-3-642-13205-6_9 doi: 10.1007/978-3-642-13205-6_9

Stock, J. H. & Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models* (pp. 80–108). Cambridge University Press.

The Football Association. (2024). *Rules of the FA Challenge Cup 2024-25.* https://www.thefa.com/-/media/thefacom-new/files/competitions/2024-25/rules/the-emirates-fa-cup-rules---season-2024-25.ashx. ([Accessed: 2024-10-12])

The Guardian. (2022). Jürgen klopp hits out at liverpool's crowded fixture schedule. *The Guardian*. (Available at: https://www.theguardian.com/football/2022/apr/12/jurgen-klopp-hits-out-at-liverpools-crowded-fixture-schedule (Accessed: 2024-02-29))

UEFA. (2025). *Country coefficients 2025.* https://www.uefa.com/nationalassociations/uefarankings/country/?year=2025. (Accessed: 2024-07-23)

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press.

Wunderlich, F., Seck, A. & Memmert, D. (2023). Skill or luck? analysing random influences on goal scoring in football. In *13th world congress of performance analysis of sport and 13th international symposium on computer science in sport* (pp. 126–129). Cham: Springer Nature Switzerland.