Peer Effects in Sports: Does Playing in Elite Tournaments Affect Performance?

İlhan Güner   Mehdi Hamidi Sahneh*

This paper asks whether participation in elite international football tournaments help teams to perform better in their domestic tournaments. We create a novel dataset by combining end of the season league tables with match level data on betting odds and match scores. Using a regression discontinuity design on top 5 European football leagues, we find that participation in highly selective UEFA Champions League (UCL) generates large performance gains to participating teams. More precisely, participation in the UCL improves i) goal differences (goals scored minus goals conceded) by approximately 0.3 goals per game and ii) probability margin (probability of winning minus probability of losing) by approximately 10 percentage points. By ruling out other potential channels, we argue that our results suggest the importance of peer effects in sports.

Keywords: Peer effects; Sport economics; Regression discontinuity design; Football tournaments.

---

*İlhan Güner: School of Economics, University of Kent, Canterbury, Kent, CT2 7NP, UK (e-mail: i.guner@kent.ac.uk.); Mehdi Hamidi Sahneh, School of Economics, University of Kent, Canterbury, Kent, CT2 7NP, UK (e-mail: mh766@kent.ac.uk.). We are grateful for helpful comments from Amrit Amirapu, Irma Clots-Figueras, Guy Tchuente, and seminar participants at the University of Kent. Address correspondence to Ilhan Güner.

1 Introduction

Economists and other social scientists have long sought to understand how ones performance is influenced by the performance of others, a mechanism referred to as "peer effect".[1] Elite sport tournaments provide an environment to study the peer effects because they sharply increase athletes' exposure to elite peers. In this paper, we use a unique game-level dataset from top 5 European football (soccer) leagues and a regression discontinuity design to study how participation in an elite football tournament affects teams' performance in their domestic national leagues.

The Union of European Football Associations (UEFA) organizes the UEFA Champions League (UCL), which is considered to be the crown jewel of the club football. Each season UCL brings together the best teams from across Europe in a highly competitive tournament.[2] Only the best teams of national leagues are eligible to participate in the UCL. UEFA allocates quotas to national leagues for the number of teams eligible to play in the UCL. These quotas generate points based eligibility cutoffs, which we exploit to measure the *causal* impact of the UCL participation on team's performance. In our analysis, we focus on teams playing in the top 5 European leagues, i.e., English Premier League (EPL) from England, La Liga from Spain, Bundesliga from Germany, Serie A from Italy, and Ligue 1 from France, from 2000 through 2019.

There are at least two reasons that the UCL participation might affect performance. First, participation in the UCL is associated with sharp increases in peer quality. Taking player valuations as a proxy for quality, we show that the average quality of players in

[1] Among others, Falk and Ichino (2006), Carrell, Fullerton, and West (2009), Mas and Moretti (2009), Gould and Winter (2009), Cornelissen, Dustmann, and Schönberg (2017), Azoulay, Graff Zivin, and Wang (2010), Bandiera, Barankay, and Rasul (2010), and Akcigit, Caicedo, Miguelez, Stantcheva, and Sterzi (2018) find empirical evidence of peer effects. Abdulkadiroğlu, Angrist, and Pathak (2014) and Guryan, Kroft, and Notowidigdo (2009) find no evidence of peer effects. Angrist (2014) discusses econometric challenges.

[2] Notice that the UCL does not replace national leagues. Teams participating in the UCL keep competing in their national leagues.

the UCL is dramatically higher than the average quality of players in national leagues. For instance, average value of players in the UCL in 2019 was about £14 million (2015 prices), £4 million higher than the average valuation in EPL, £7 million higher than the average valuation in La Liga, much higher than average valuation of players in other leagues. Thus, the UCL provides a setting in which participants might learn from their elite peers or be motivated by them. Second, the UCL participation is associated with significant financial rewards. For instance, according to UEFA, the 32 clubs that played in the 2018/19 UCL group stage have shared about €2 billion in payments from UEFA. Clubs might use these financial resources to improve their performance by signing better payers and managers.[3]

The analysis proceeds in three stages. We first use a variety of sources to construct a novel dataset, which encompasses the universe of football matches across top 5 European football leagues. We collect information on betting odds, match scores, and end of the season points and rankings. We combine such information with a complete list of player transfer fees and team manager changes. We use the data to construct two measures of performance of a team at the match level: the goal difference, defined as the number of goals scored minus the goal conceded, and probability margin of winning, defined as the (betting market assigned) probability of winning minus the probability of losing.

In the second stage, we investigate whether participation in the UCL generates performance gains for teams in their domestic leagues. To answer this question, we use a fuzzy regression discontinuity design that compares performance among teams that narrowly qualified to participate in the UCL and teams that narrowly lost the opportunity. We find that participation in the UCL increases the probability of winning a game by about 8 percentage points, and the within game goal difference by about 0.3 goals. These estimates are statistically significant and robust across different specifications.

[3]By manager we mean a person who is in charge of training and performance of a team, not an executive manager.

In the third stage, we investigate the causal channels through which participation in the UCL might have affected team performance. Because playing in the UCL is associated with a sharp increase in peer quality, "peer effect" is one plausible channel. However, there are other channels that could be driving the performance. For instance, participants may use the financial reward from the UCL (or other revenue sources) to change their team composition (e.g., by hiring better players and managers), which might affect their relative performance. To rule out changes in the team composition as a causal channel in our results, we use the same regression discontinuity idea to investigate team composition near the eligibility cutoff. First, we show that transfer fees close to the UCL cutoff are similar, so it is not the case that the teams that narrowly qualify spend significantly more money to sign better players, compared to the teams that narrowly miss the opportunity. Second, the improvement in team performance persists even when we restrict the sample to the teams with no managerial changes before the start of the season, so managerial changes cannot explain the jump in team performance at the eligibility cutoff. These evidence thus rules out team composition as an explanation of our findings and suggest that peer effects contributes to the team performance in ways that are hard to reconcile with team composition.

Our paper contributes to a large literature on peer effects. Perhaps surprisingly, one of the first studies in this literature was conducted in a sport setting. Triplett (1898) observed that cyclists ride faster when competing with other cyclist, compared to when they race alone or against a pace-maker, and concluded that the presence of others affects performance. Consequently, many studies examined whether and how one's performance is influenced by the performance of others in various settings. Examples include education (Carrell, Fullerton, and West, 2009, Sacerdote, 2001), controlled laboratory experiment (Falk and Ichino, 2006), workers in the workplace (Mas and Moretti, 2009, Bandiera, Barankay, and Rasul, 2010), and scientists (Azoulay, Graff Zivin, and Wang, 2010, Waldinger, 2012), among many others.

In this direction, elite sport tournaments are an important causal channel to explore because they sharply increase exposure to elite peers. The most related paper to ours is Guryan, Kroft, and Notowidigdo (2009), who find no evidence that participation in professional golf tournaments affects performance. In contrast we find robust evidence of performance gains from participation in the UCL. As noted by Ehrenberg and Bognanno (1990), golf tournaments, or other settings in which payments depend only on relative rank compared to other competitors, the payment structure influence player's performance. An advantage of our setting is the clear distinction between payments from participation in the UCL and domestic performance. This distinction makes it easier to disentangle whether the observed performance gain is due to peer effects or financial incentives. If the monetary incentives from the UCL were correlated to the performance in the domestic leagues, this would confound the identification of peer effects, as improved performance in the domestic league could be stemming from responding to financial incentives.

Identifying peer effects is challenging for other reasons as well. The standard approach in this literature consists of estimating an outcome-on-outcome specification. However, as Angrist (2014) points out, outcome-on-outcome regressions are likely to produce biased estimates, with both the sign and the size of the bias depend on the true underlying data generating process. Several studies attempt to solve these problems by exploiting quasi-experimental variation that comes close to the ideal experiment. Closely related to our paper is the Abdulkadiroğlu, Angrist, and Pathak (2014) and Zimmerman (2019) who exploit the regression discontinuity in the selective school admissions on academic performance and social mobility. Similar to selective schools, participation in the UCL is associated with sharp increases in peer quality, which we exploit to investigate peer effects. As Angrist (2014) argues, this strategy is well suited to identify peer effects; however, it comes at the cost of requiring further investigation of plausible channels through which treatment may affect performance.

The rest of the paper proceeds as follows. Section 2 provides more details about our research design and the UEFA Champions League, and Section 3 describes the data used in this analysis. Section 4 outlines the empirical strategy and its application to the analysis of the UCL program. Section 5 reports the relevant identification checks. Section 6 shows and discusses the main results and some extensions. Section 7 concludes.

2 Institutional Background

European football (soccer) is structured around national football associations. Each national football association organizes (or oversees) many hierarchical divisions of football leagues. At the end of each season, the top ranked teams in a division are promoted to the next upper division, whereas the lowest ranked teams are relegated to the next lower division. Throughout this paper, we will focus only on the top divisions from England (EPL), Spain (La Liga), Germany (Bundesliga), Italy (Serie A), and France (Ligue 1), and will refer to these top divisions as *national leagues*. These football leagues are mostly regarded as the top 5 football leagues in Europe. In fact, 39 out of 40 finalists of the UCL in the last 20 years are from these leagues.

Union of European Football Associations (UEFA) is an umbrella organization of national football associations. Besides overseeing national football associations, UEFA organizes two big club competitions: UEFA Champions League (UCL) and UEFA Europa League (UEL). The UCL is the most prestigious club competition in European football, contested by 32 clubs from the strongest UEFA members. Participating teams play both in UEFA competitions and in their national leagues. Due to the prestige and financial incentives of this tournament, every club wants to play in the UCL. In its present format, less than 20% of teams from each national leagues are *eligible* to play in the UCL.

Eligibility is mostly determined by the team's performance in its national league. Each national league is contested by N teams, playing twice (i.e., home and away) against each opponent. The result of each match is decided by the goal difference, defined as goals scored minus goals conceded. A positive goal difference within a match indicates a win, a zero goal difference indicates a draw, and a negative goal difference indicates a loss. Accordingly, in each match a team earns 3 points for a win, 1 point for a draw, 0 points for a loss. The ranking of the teams at the end of the season is a deterministic function of the total number of points collected by each team during that season. In case two teams have the same number of points then the better placed team will be the team with better total goal difference,[4] or better goal difference in direct games amongst the tied teams.

Eligibility is determined at the end of each season, with the winner and 2-3 runner-ups are eligible for playing in the UCL the next season. The number of teams from each member association entering into the UCL is based on the UEFA coefficients of the member associations.[5] The higher an association's coefficient, the more teams represent the association in the Champions League. In reality, however, eligibility does not necessarily imply playing in the UCL.[6] More precisely, the UEFA coefficient indicates the number of teams directly play in the UCL, and the number of teams that must go through playoff-rounds, with some small chance of elimination.[7] For instance, a total of 4 teams out of 20 teams in the English Premier League (EPL) were eligible for the 2015-16 UCL season, with the top 3 teams from 2014-15 EPL final table qualified automatically, but the 4th team went to a playoff-round.

Teams who qualify to play in the UCL see a dramatic change in their peer quality compared to the teams that narrowly miss this opportunity.[8] Figure 1 shows average

[4] Total goal difference is calculated as the sum of within game goal differences in a given season.

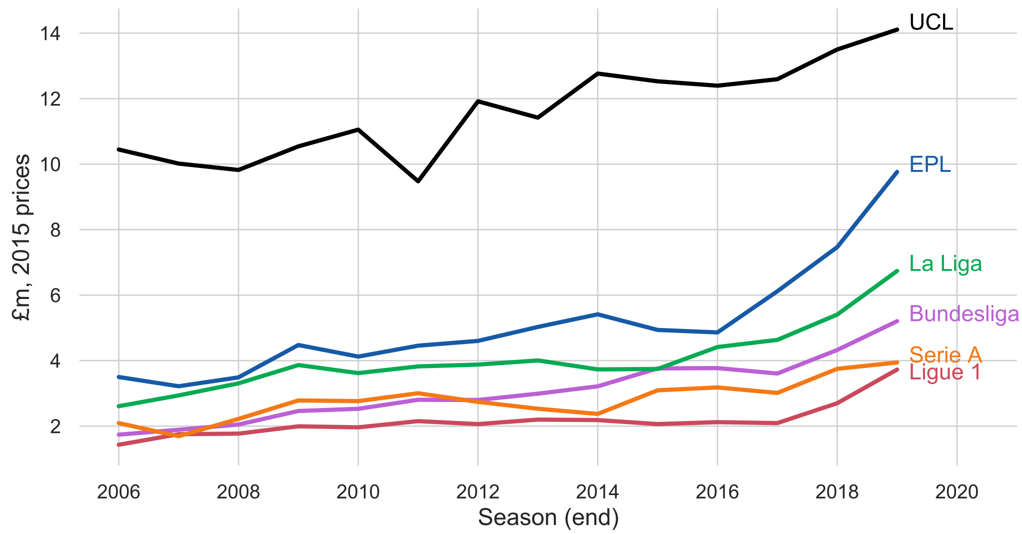
[5] UEFA calculates these coefficients based on the results of clubs representing each association during the previous five Champions League and UEFA Europa League seasons. See `https://www.uefa.com/memberassociations/uefarankings/club` for more information.

[6] Appendix table (5) provides more information on the UCL eligibility and participation.

[7] It is a small chance because the teams from top 5 European countries usually face teams from countries with lower UEFA coefficients (i.e., weaker). See UEFA Article 3 for more information about entries for the competition.

[8] Teams that narrowly miss the opportunity to play in the UCL, most likely participate in the less prominent Europa League (UEL). In our analysis, we only focus on the UCL cutoffs since the eligibility for the Europa league was partly based on applications, and not role-based cutoffs. Moreover, clubs that are knocked out of the qualifying round and the group stage of the Champions League join the Europa League, at different stages, which makes the peer definition more difficult.

Figure 1: Average Market Value of Players



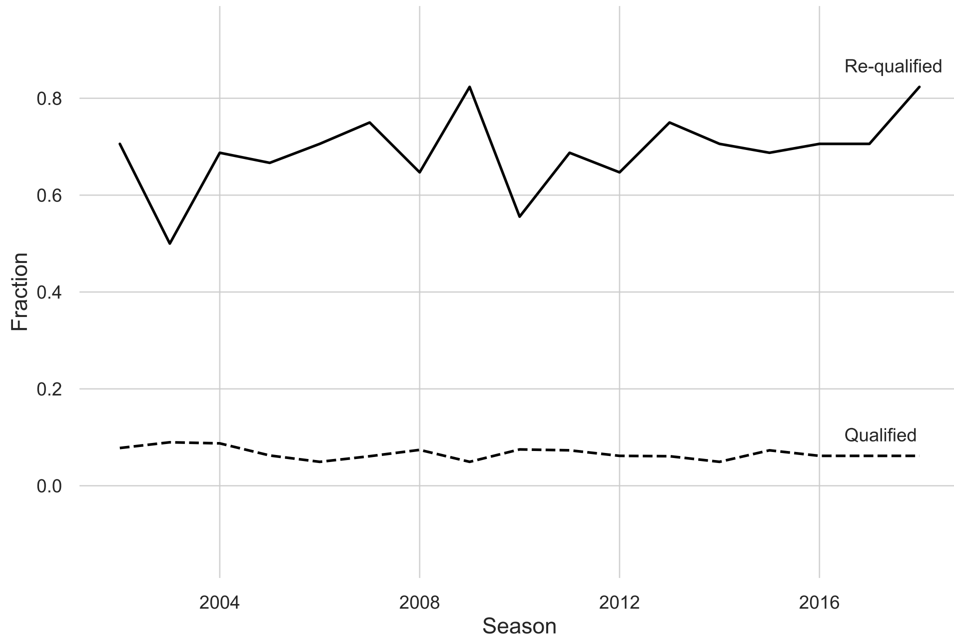
Notes: Unweighted means of valuations of players registered in each league. UCL average includes players from all the participating teams, not just players from the top 5 European national leagues.

player valuations in 5 European football leagues and in the UCL.[9] We see that average valuation of players in the UCL is consistently higher than average valuation of players in other leagues. This is not surprising since the UCL is a competition in which only elite teams with top players can participate.

One of the striking facts of European football is the consistently high rate of success among the top teams. Figure 2 shows that the teams that participate in the current UCL season have about 70% chance to participate in the next UCL season. By contrast, the teams that do not participate in the current UCL season have less than 10% chance to participate in the next UCL season. These rates have been fairly stable over the last two decades. Therefore, higher ranked teams tend to do well in the next season and participate in the UCL in the following season. Such persistence in the UCL participation can be a result of two factors: i) participating teams are inherently better than others and ii)

[9] Player valuations are taken from `https://www.transfermarkt.co.uk`. Average player value is mean value of all the registered players in a league.

Figure 2: Persistence of Participation in UCL



Notes: This figure plots the UCL participation rates in season$t+1$, conditional on participation status in season t. Solid line show the re-qualification rate, i.e., teams that participate at both season t and $t+1$, while the dashed lines show the fraction of teams that play in the UCL in seaosn $t+1$, but did not play in the UCL in season t. The sample used to construct this figure consists of teams from top 5 European national leagues from 2000 to 2019.

by participating in the UCL, teams improve their domestic league performance. In this study, we examine whether and how the second factor might have played a role.

3 Data Description

To answer our research questions, we compile a new data set of European football at the match level. Our dataset contains information on the universe of football matches across top 5 European Football leagues, namely EPL (England), La Liga (Spain), Bundesliga (Germany), Serie A (Italy), and Ligue 1 (France), from 2000 to 2019. We collected information on betting odds and match scores (goals scored and goals conceded) from `https://www.football-data.co.uk` Web site.

We construct two measures of a team's performance in a match. First measure is an ex-post measure, *goal difference* within a match, which is the number of goals scored by a team minus the number of goals conceded by the team in the same match. Consider a match between two teams, say *Home vs Away*. If *Home* were to score six goals but concede one goal (i.e., *Away* scores one goal), the goal difference would be +5. More specifically, we define the goal difference as

$$\mathrm{GD}_{i,j,h,l,t} = \mathrm{GS}_{i,j,h,l,t} - \mathrm{GS}_{j,i,h,l,t},$$

where $\mathrm{GS}_{i,j,h,l,t}$ denotes the goals that team i scores against team j, $h \in \{0, 1\}$ indicates whether the game is played at home or away, l is the league and t is the season.

However, goal difference depends partly on the dynamics of the match, and might not entirely reflect team performance. To use the *Home vs Away* example above, suppose that *Home* is significantly stronger than *Away*, and would normally win the match with +5 goal difference if they exert full effort. However, after achieving a comfortable lead over *Away* (e.g., +3 goals), *Home* might reduce their efforts to preserve energy for the next match. Our assumption here is that effort improves goal difference, but effort is costly. Thus, while we would expect that *Home* to win, we do not expect the goal difference to precisely reflect the difference in quality of the two teams.

The second measure of team performance is an ex-ante measure, *probability margin* for a match: a team's market determined winning probability minus its opponent's winning probability before the start of the match. To construct probability margins, we obtain match odds from 13 major online bookmakers: Bet365, Blue Square, BWin, Gamebookers, Interwetten, Ladbrokes, Pinnacle, Sporting Odds, Sportingbet, Stan James, Stanleybet, VC Bet, and William Hill.[10] The data is obtained from `https://www.football-data.`

[10]For each match, we have betting data from most of these bookmakers, but not necessarily from all bookmakers.

co.uk, which is unique with respect to its size and the information it contains: The dataset spans the period 2000-2019 containing information about 27461 unique football matches. For our purposes, the variables of interests are: draw, home win, and away win odds. The odds are kick-off time odds (also known as *closing odds*), i.e., those that were quoted when bookmakers stopped accepting new bets before the matches.

The odds represent the current balance of opinions about the likelihood of a team winning as expressed by the amounts of money wagered for and against it. To fix ideas, let's think of a game between two hypothetical teams: *Home* vs. *Away*. Typically, bookmakers determine their odds based on a statistical model, which takes all the available information into consideration, including the team's lineup, injuries, location (home or away), current form and historical performance. Once the initial odds have been set it will be adjusted based on the amount of money put on the different outcomes by traders. If a bookmaker underpriced the odds of a particular outcome, let's say *Home* win, then traders will put money on this outcome until it is priced at a fair value. For instance, if a trader places, say, $100 on the *Home* to win, the odds will shift. If another trader believes that the odds are now mispriced and that there is value on the other side, they might place $100 on the *Away* to win and the odds will shift again and thus eliminating the mispricing.

Typically, sports betting odds are expressed as decimal odds.[11] Decimal odds describe the total return, including both stake and profit, if the bet wins. For instance, odds of 1.25 would imply that a $100 winning stake will return $125 in total (including the original stake of $100).[12] Consequently, we can obtain the implied (observed) probabilities from

[11] See Buchdahl (2016) for more information on betting odds.

[12] Fractional odds simply describe the potential profit that can be won from a unit stake. Consequently, odds of 1/4 (one-to-four) would imply that the bettor with a winning stake of $100 will make a profit of $25. It is straightforward to convert fractional odds into decimal odds, using the equation

$$\text{Decimal odds} = \text{Fractional odds} + 1$$

decimal odds, using the equation

$$\text{Implied probability} = \frac{1}{\text{Odds}}$$

For example, a home-draw-away book with odds of $\mathrm{O}_h = 1.53$, $\mathrm{O}_d = 3.5$, and $\mathrm{O}_a = 5.5$ implies probabilities

$$\mathrm{P}_h = \frac{1}{\mathrm{O}_h} = 0.654,\ \mathrm{P}_d = \frac{1}{\mathrm{O}_d} = 0.286,\ \mathrm{P}_a = \frac{1}{\mathrm{O}_a} = 0.182$$

where O_h, O_d, and O_a are the home team win, draw, and away team win odds and P_h, P_d, and P_a denotes the home team win, draw, and away team implied probabilities.

These probabilities, however, do not reflect the "fair" odds.[13] More precisely, the sum of the probabilities exceeds 1, and equals 1.121 in the above example. Mathematically, of course, the sum of probabilities for all possibilities must be 1. The excess 0.121 in our example determines the bookmaker's profit margin. Thus, the bookmaker's odds do not reflect the fair (true) probabilities. To obtain the fair probabilities, we first need to remove the margins that bookmakers apply to their odds. Since bookmakers usually do not reveal how they apply the margins to their odds, we are forced to guess how they might do it. The common method to obtain the fair odds is to assume that the margin applied to each outcome is proportional to the outcome probability.[14] Thus, the fair probability for the i-th outcome, P_i, is

$$\mathrm{P}_i^* = \frac{\mathrm{P}_i}{\sum_i \mathrm{P}_i},\ i \in \{h, d, a\}.$$

To use the example above, the fair probabilities the home team win, draw, and away team

[13]If the odds are equal to the true odds that an event will occur, then they are said to be "fair" odds.

[14]Our results are both qualitatively and quantitatively very similar when we use other methods (e.g., additive method or logarithmic method) to remove the markups.

win odds are given by

$$\mathrm{P}_h^* = \frac{0.654}{1.121} = 0.58,\ \mathrm{P}_d^* = \frac{0.286}{1.121} = 0.25,\ \mathrm{P}_a^* = \frac{0.182}{1.121} = 0.16$$

To calculate our ex-ante measure of team performance, probability margin, we use fair probabilities of home team and away team winning for each match and from each bookmaker. More specifically, the probability margin of home team against away team is calculated as follows:

$$\mathrm{PM}_{i,j,h,l,t} = \mathrm{P}^*_{i,j,h,l,t} - \mathrm{P}^*_{j,i,h,l,t},$$

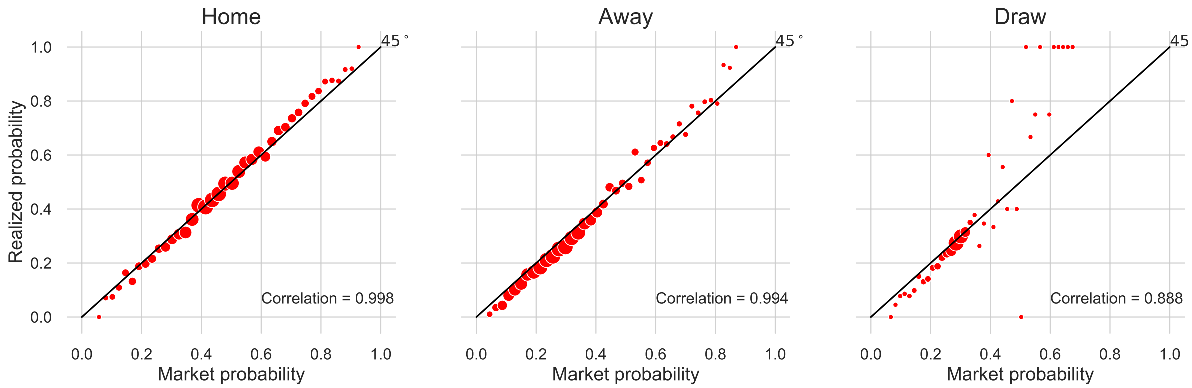
where $\mathrm{P}^*_{i,j,h,l,t}$ denotes the fair probability that team i wins against team j, $h \in \{0, 1\}$ indicates whether the game is played at home or away, in league l and season t.

For the purpose of our analysis, probability margins from all bookmakers with available data have been combined into a single probability by taking cross-sectional average of probability margins over bookmakers.[15] For a few matches, we don't have betting odds from any bookmaker. When no betting information is available, we remove that observation from our sample even if we have information about the goal difference. Notice that for each match we construct two probability margins, one for the home team and one for the away team.

The betting markets are doing a remarkable job at predicting actual results. To see this, we compare the market predictions with the actual outcomes in Figure 3. Remember that for each match we have home team winning, away team winning, and draw probabilities constructed from corresponding odds. For each probability, we split matches into 40 bins with a bin size equal to 1.6 percentage points. Then we calculate the ratio of the matches in each bin that is in accordance of the market prediction. For instance, we take all games for which the market predicts that the probability of the home team beating the away team is between 5.4 percent to 7 percent. We then report the proportion of

[15] Bürgi and Sinclair (2017) show that it is difficult to improve upon the simple cross-sectional average.

Figure 3: Market Prediction and Actual Results



Notes: Market Probability refers to the probability that the market predicts a team will beat the other team (Home or Away team winning probabilities) or the game will result in a draw, and Realized Probability refers to actual success rates in the sample. The sample used to construct this figure consists of all games from top 5 European countries from 2000 to 2019. The dots in the figure are averages of the probabilities of different events calculated in bins 1.6 points wide, while the line through the dots shows a perfect fit. The size of a dot is proportional to the number of matches in the bin corresponding to the dot.

the games where the home team beats the away team. From Figure 3 it is clear that the market probabilities correspond quite closely to the actual results: When the market predicts that the probability of home team win is 5.4-7 percent, the home team wins about 5.4-7 percent of the time. It is only for the case of draws that the market and the actual probabilities are not closely aligned. In this case, however, the sample size is relatively small as represented by the sizes of the dots in the figure; there are relatively few games which the market predicts to be a draw with probability greater than 50 percent.

Our analysis combines match level data with team level data from various sources. League tables are collected from Wikipedia and provide end of the season information about total points earned and ranking of each team. UCL quotas are collected from Wikipedia and specify the number of teams from each league at each season that can directly play in the UCL group stage and the number of teams that need to play playoff-rounds to qualify for the UCL group stage.[16] Finally, we obtain the data on transfer fees and managerial changes from `https://www.transfermarkt.com`.

Table 1 shows descriptive statistics on our outcome variables in the entire sample and in each league separately. Distributions of goal difference and probability margin are similar across leagues. Mean probability margin in each league is close the mean probability margin in the entire dataset, 0.03. Mean goal difference in each league is also close to the mean goal difference in the entire sample, 0.08. Due to the discrete nature of goal difference, the percentiles recorded in Table 9 are identical across leagues. Distribution of transfer fees differ across leagues with EPL teams spending considerably higher per each signing than the other leagues. League 1 teams, on the other hand, receive more transfer fees than they pay.

4 Identification Strategy

We are interested in whether and how participation in the UCL might have affected teams performance. Any strategy that aims at identifying such causal effects needs to address the endogeneity in the UCL participation status. In practice, football teams differ along many dimensions, and certain teams may be more likely to participate in the UCL (e.g., those with better organizational structure). Our empirical approach overcomes the endogeneity problem by focusing on the jump in performance among teams at the eligibility threshold. We do this using a fuzzy regression discontinuity design that compares performance among teams that narrowly played in the UCL and narrowly did not.

Our econometric strategy therefore begins by constructing a running variable that determines treatment assignment. As we discussed in Section 2, eligibility depends on the ranking at the end of the season, which itself is a function of total points. Thus, we

[16] We classify teams that played at least in the group stage. Teams knocked out during the play-off rounds are not classified as the UCL participants.

Table 1: Descriptive Statistics

				Percentile				
	Obs	Mean	SD	10th	25th	50th	75th	90th
				All leagues				
GD(t+1)	54922	0.08	1.78	-2.00	-1.00	0.00	1.00	2.00
PM(t+1)	54922	0.03	0.34	-0.42	-0.20	0.04	0.28	0.49
Fees(t+1)	13739	0.66	10.18	-6.48	-1.51	0.32	2.84	8.22
				Bundesliga				
GD(t+1)	9452	0.07	1.91	-2.00	-1.00	0.00	1.00	2.00
PM(t+1)	9452	0.03	0.33	-0.41	-0.20	0.03	0.26	0.47
Fees(t+1)	2797	0.33	6.52	-3.38	-0.51	0.19	1.70	4.88
				EPL				
GD(t+1)	11362	0.10	1.82	-2.00	-1.00	0.00	1.00	2.00
PM(t+1)	11362	0.04	0.36	-0.46	-0.22	0.05	0.31	0.53
Fees(t+1)	2963	2.16	12.51	-7.38	-1.97	0.83	5.76	14.04
				La Liga				
GD(t+1)	11626	0.08	1.84	-2.00	-1.00	0.00	1.00	2.00
PM(t+1)	11626	0.03	0.36	-0.45	-0.21	0.03	0.28	0.51
Fees(t+1)	1989	0.73	13.10	-8.46	-2.48	0.45	3.40	10.19
				Ligue 1				
GD(t+1)	11491	0.07	1.67	-2.00	-1.00	0.00	1.00	2.00
PM(t+1)	11491	0.03	0.30	-0.36	-0.18	0.02	0.24	0.41
Fees(t+1)	2168	-0.22	9.37	-7.14	-2.16	0.36	2.21	5.52
				Serie A				
GD(t+1)	10991	0.09	1.67	-2.00	-1.00	0.00	1.0	2.00
PM(t+1)	10991	0.04	0.35	-0.44	-0.21	0.05	0.3	0.51
Fees(t+1)	3822	0.21	8.86	-6.08	-1.44	0.27	2.5	6.81

Notes: This table shows descriptive statistics on the main outcome variables used in the analysis for the entire sample and for each league in the sample. Positive values for (transfer) Fees represent incoming transfers, whereas negative values represent outgoing transfers.

construct our running variable as a function of total points.[17] More precisely, we first calculate the league-season-specific cutoff point as the average of total points of the worst eligible teams and best ineligible team. For instance, if 4 teams from league l are eligible to participate in the UCL in season $t+1$, the running variable is defined as

$$\text{Pts}^{*}_{l,t} = \frac{\text{Pts}^{4th}_{l,t} + \text{Pts}^{5th}_{l,t}}{2}$$

where $\text{Pts}^{4th}_{l,t}$ and $\text{Pts}^{5th}_{l,t}$ denotes the total points of the $4th$ team the $5th$ team from league l in season t, respectively.

To account for the difference in cutoff points across leagues and seasons, we center and scale the running variable around the cutoff value. Precisely, our standardized running variable is then defined as

$$\text{S}_{i,l,t} = \frac{\text{Pts}_{i,l,t} - \text{Pts}^{*}_{l,t}}{\text{Std}(\text{Pts}_{l,t})},$$

where $\text{Pts}_{i,l,t}$ denote team i's points, $\text{Pts}^{*}_{l,t}$ is the league-season-specific cutoff, and $\text{Std}(\text{Pts}_{l,t})$ is the standard deviation of the league-season total points. These standardized league-season-specific points equal zero at the cutoff, with nonnegative values indicating teams who are eligible to play in the UCL in the next season. Thus, the eligibility is a deterministic function of the standardized points

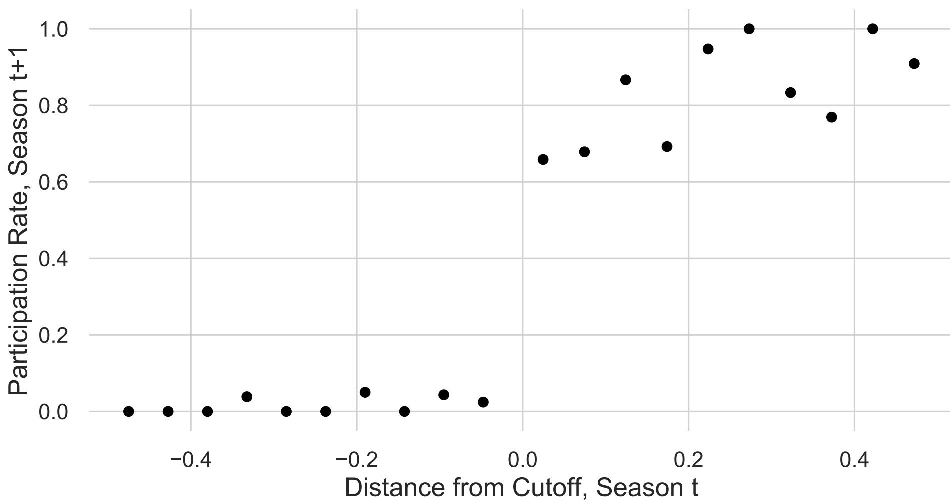
$$\text{Elig}_{i,l,t} = \mathbb{1}(\text{S}_{i,l,t} \geq 0),$$

which assigns all teams whose score are below the zero cutoff to the control group, and all teams whose score is above zero to the treatment group.

Although eligibility is a deterministic function of standardized points, participation in

[17] We don't define our running variable based on team ranks because proximity in ranking does not imply proximity in performance before treatment. Intuitively, by defining the running variable according to team ranks, we might compare $4th$ team against the $5th$, while they are very different from each other based on their total points.

Figure 4: Discontinuity in Probability of Participation in the UCL



Notes: This figure plots participation in the UCL group stage, plotted against league-season-specific standardized running variable.

the UCL remains probabilistic. Specifically, not all eligible teams play in the UCL. Figure 4 gives a graphical representation of the participation rate in the UCL as a function of the standardized running variable. The plot clearly shows that less than 5% of teams that were not eligible, participated in the UCL. The ineligible teams that participate in the UCL are the UEFA Champions League and UEFA Europa League titleholders. For instance, Liverpool FC ranked 5th in the 2004-05 EPL season, so not eligible based on standardized points, but actually played in the 2005-06 UCL season since they won the UCL in the 2004-05 season. Above the threshold, the probability of participating in the UCL increases rapidly: More than 70% of the eligible teams participate in the UCL. Teams that were eligible but did not play in the UCL were those teams that lost the playoff rounds.

This setup naturally leads to a fuzzy RD design, where standardized points ($\mathrm{S}_{i,l,t}$) is the running variable that partially determines participation in the UCL. As discussed in Hahn, Todd, and Van der Klaauw (2001), estimation of the UCL treatment essentially amounts to a simple 2SLS estimation strategy, using the discontinuity in the eligibility as an instrumental variable for the UCL participation status. More precisely, let $\mathrm{Y}_{i,j,h,l,t+1}$ is an outcome variable of team i against team j, $h \in \{0,1\}$ indicates whether the game is played at home or away, from league l in season$t+1$. To obtain the causal impact of the UCL participation, we estimate variants of the following parametric regression model:

$$\mathrm{Y}_{i,j,h,l,t+1} = \alpha + \tau \mathrm{UCL}_{i,l,t+1} + f(\mathrm{S}_{i,l,t}) + \lambda_t + \Lambda_l + \epsilon_{i,j,h,l,t+1}, \tag{1}$$

where $\mathrm{UCL}_{i,l,t+1}$ is the indicator for participation in the UCL (i.e., treatment status), and $f(\mathrm{S}_{i,l,t})$ is a flexible function of the vote margin, that is allowed to differ on each side of the discontinuity. We follow the common practice in the literature, and assume that $f(.)$ can be described by a low-order polynomial.[18] Furthermore, λ_t and Λ_l are dummy variables capturing leagues and seasons fixed effects, respectively, and $\epsilon_{i,j,h,l,t+1}$ is an error term.

The parameter of interest is τ, which captures the casual impact of participation in the UCL. A consistent estimate of τ can be obtained by estimating (1) with the instrumental variable estimator, where $\mathrm{Elig}_{i,l,t} = \mathbb{1}(\mathrm{S}_{i,l,t} \geq 0)$ is used as instrument. The corresponding first-stage in this case is

$$\mathrm{UCL}_{i,l,t+1} = \gamma_0 + \gamma_1 \mathrm{ELig}_{i,l,t+1} + f(\mathrm{S}_{i,l,t}) + \lambda_t + \Lambda_l + \nu_{i,l,t+1}, \tag{2}$$

where the dummy variable $\mathrm{Elig}_{i,l,t+1}$ is used as an instrument for $\mathrm{UCL}_{i,l,t+1}$.

5 RD Validity Checks

Before presenting our results, we conduct several checks to ensure the validity of our RD strategy. The key identifying assumption in our RD design is the inability of teams to precisely control treatment status. Density tests, first proposed by McCrary (2008), seeks to formally determine whether there is evidence of manipulation of the running variable at

[18] Gelman and Imbens (2019) argue that including high-order polynomials of the running variable may lead to noisy estimates and poor coverage of confidence intervals.

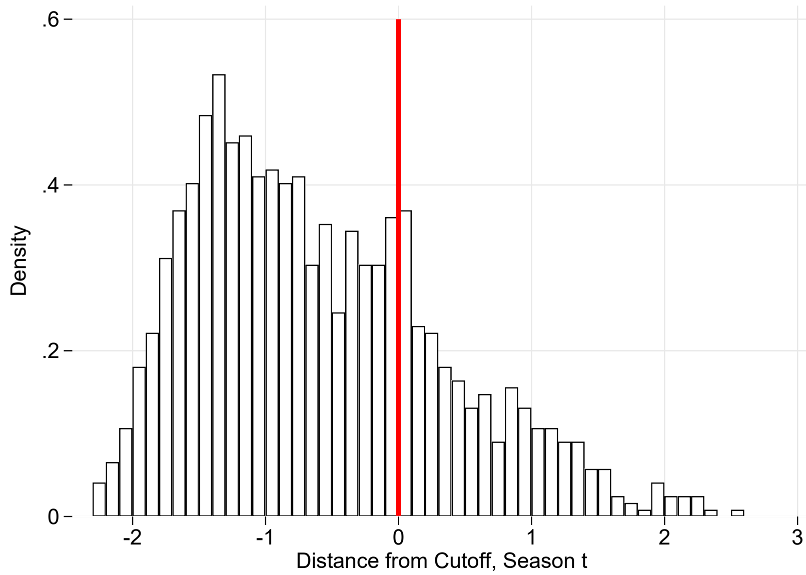
the cutoff. In our case, local random assignment would be violated if teams just below the cutoff could influence their total number of points to be eligible for the UCL in the next year. Violation of local random assignment requires some teams to be able to precisely control the outcome of the games they play against their opponents. However, in our setting, points are gained (and lost) against direct opponents, who also want to rank as high as possible in their domestic league and play in the UCL. Furthermore, there is also some element of chance involved in a match outcome, which can influence the total number of points and eligibility at the end of the season. This supports our RD design from the outset, since it is unlikely that some teams could precisely control the assignment variable.[19]

To further ensure that our results are reliable, we test the null hypothesis that the density of the running variable is continuous at the cutoff. Inspecting the density of the running variable, shown in Figure 5, suggests no manipulation of the assignment variable. Furthermore, the Cattaneo, Jansson, and Ma (2019) density test confirms this point. The test statistic is 0.303 (p-value is 0.762), and therefore, we fail to reject the null hypothesis of no difference in the density of running variable at the cutoff.

As a second validity test, we check the continuity of the predetermined variables across the threshold. Lee (2008) argues that if predetermined variables are unaffected by the treatment, the balance of predetermined variables should be smooth through the threshold. In our context, the key practical issue is that our predetermined variables might be actually treated, because some teams that participate in the UCL season $t+1$ might have also participated in the UCL in season $t-1$. For this reason, when analyzing the balance of the predetermined variables, we restrict the sample to the teams that did not play in the UCL in season $t-1$.

[19] The 2006 Italian football scandal, or Calciopoli, where a number teams tried to influence referee appointments, is a valid concern. In our analysis, we drop observations from Italian Serie A for season 2006, but the result are qualitatively and quantitatively are similar when we include these observations in our sample.

Figure 5: Density of Running Variable



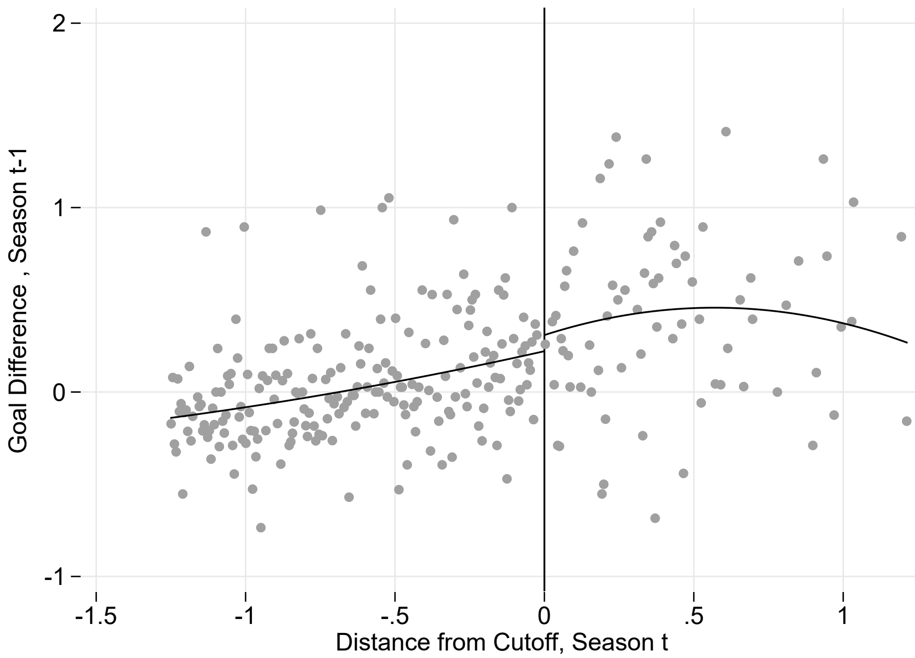
Notes: Density of the standardized points within bins of width 0.10.

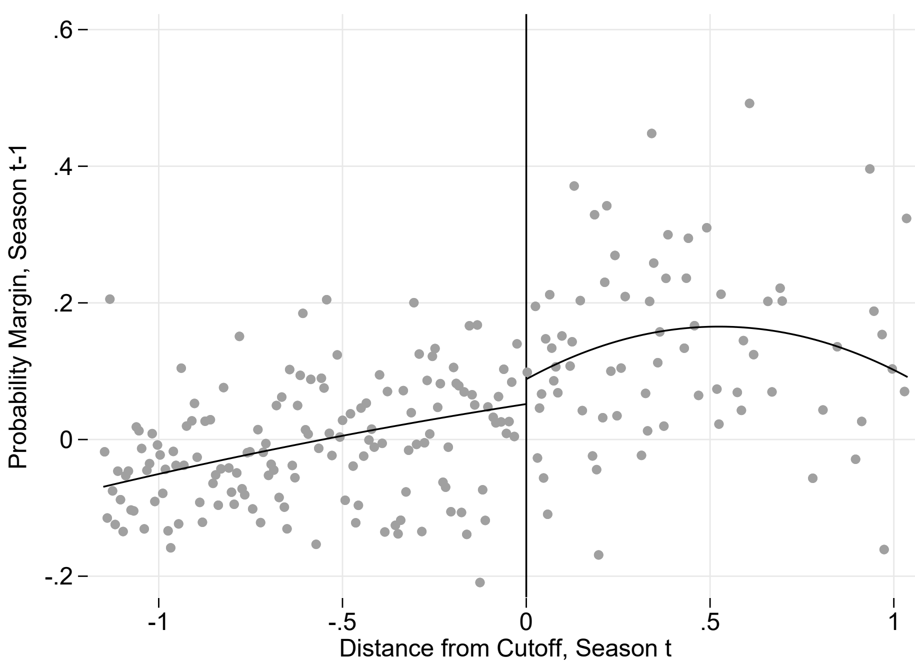
Figure 6 plots the balance of the predetermined variables close to the cutoff. Each dot in the figure indicate local averages, and the lines are fitted values from a quadratic polynomial fit, which is allowed to be different on either side of the discontinuity. Throughout the paper, we use the data-driven method of Calonico, Cattaneo, and Titiunik (2015) to choose the location and the number of bins. From these RD plots, it seems more natural to assume that the underlying conditional expectations of the outcome variables given the running variable is non-linear, particularly on the treatment (right) side.[20] The bandwidth is selected optimally using the data-driven method of Calonico, Cattaneo, and Titiunik (2014). Following the recommendations of Imbens and Lemieux (2008), we report the estimates using a uniform kernel. We also experimented with different bandwidth and polynomial orders, but the results (see Appendix B) were not particularly sensitive to the bandwidth and polynomial order choice.

Figure 6 shows that eligible teams had somewhat better performance in season $t-1$.

[20]The analogous plots with linear fits are reported in Appendix B.

Figure 6: Discontinuity in Predetermined Variable





Notes: Team's goal difference (top) and probability margin of winning (bottom) in season $t-1$, by distance from the cutoff in season t. Vertical lines indicate the cutoff, and dots indicate local averages. The solid lines are predicted values from quadratic polynomial on either sides of the cutoff.

Table 2: Discontinuity Estimates in Predetermined Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Robust Bias-corrected				Conventional Method			
Dep. variable	GD(t-1)	GD(t-1)	PM(t-1)	PM(t-1)	GD(t-1)	GD(t-1)	PM(t-1)	PM(t-1)
Estimate	0.169	0.116	0.058	0.045	0.223*	0.163	0.076*	0.060
Std. Error	0.150	0.175	0.047	0.058	0.130	0.157	0.040	0.051
Bandwidth	0.792	1.254	0.791	1.153	0.792	1.254	0.791	1.153
Polynomial	1	2	1	2	1	2	1	2
Eff. Sample Size	14,097	23,766	14,097	21,610	14,097	23,766	14,097	21,610

Notes: Estimates are based on linear and quadratic polynomial on each side of the cutoff, within a MSE-optimal bandwidth and triangular kernel. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 2, columns (1)-(4), presents corresponding discontinuity estimates using the bias-corrected procedure proposed by Calonico, Cattaneo, and Titiunik (2014). The discontinuity estimates are all positive, but not statistically significant at 10% level. Given the potential for error correlation across games played by a team in a given season, we cluster standard errors two-ways, at the team-season level throughout the paper.[21] All specifications include league fixed effects and seasons fixed effects to control for differences across leagues and seasons. Table 2, columns (5)-(8), presents corresponding discontinuity estimates from conventional methods. These estimates are slightly larger in magnitude and statistically significant at 10% level for local linear specifications. However, as the Figure 6 suggests, the linear specification seems inadequate, particularly on the treatment side.

Overall, these results suggest that the predetermined variables do not vary discontinuously at the eligibility threshold.

[21] Moulton (1990) shows that in regression of individual-specific outcomes on aggregate explanatory variables, the usual standard errors are biased downwards dramatically.

6 Empirical Results

In the first part of this section, we estimate the effect of participation in the UCL on team performance in their domestic leagues. We do this using a regression discontinuity design that compares performance among teams that narrowly qualified to play in the UCL and teams that narrowly missed this opportunity. In the second part, we investigate the causal channels through which participation in UCL might have affected performance.

6.1 The Effects of the UCL Participation

Figure 7 illustrates the discontinuity in the performance of the teams right at the cutoff point. As top panel in Figure 7 clearly shows, teams who narrowly qualified to play in the UCL are much more likely to have a better goal difference in their games next season, compared to teams who narrowly did not qualify. Columns (1)-(2) of Table 3 reports estimated causal impact of playing in the UCL on the goal difference within a game, using a linear and quadratic polynomial specification. The causal effects of playing in the UCL are 0.29 goals per game in the linear specification and 0.314 goals per game in the quadratic specification. The estimates are statistically significant at 5% level, and robust to the polynomial order and bandwidth choice.

Columns (3)-(4) of Table 3 present analogous estimates for the probability margin of winning in season $t+1$. The causal effects are 0.10 in the linear specification and 0.116 in the quadratic specification, which indicates that playing in the UCL increases the winning probability by about 10 percentage points per game. These estimates are statistically significant at the 5% level, and robust to different specifications and bandwidth choice.

Columns (5)-(8) of Table 3 present analogous estimates using the conventional method. These estimates are similar in size and significance to the discontinuity estimates obtained using the bias-corrected method of Calonico, Cattaneo, and Titiunik (2014). We conduct further robustness checks, using different bandwidth and polynomial orders, but the results were not particularly sensitive to bandwidth and polynomial order choice. See Table 7 in Appendix B.

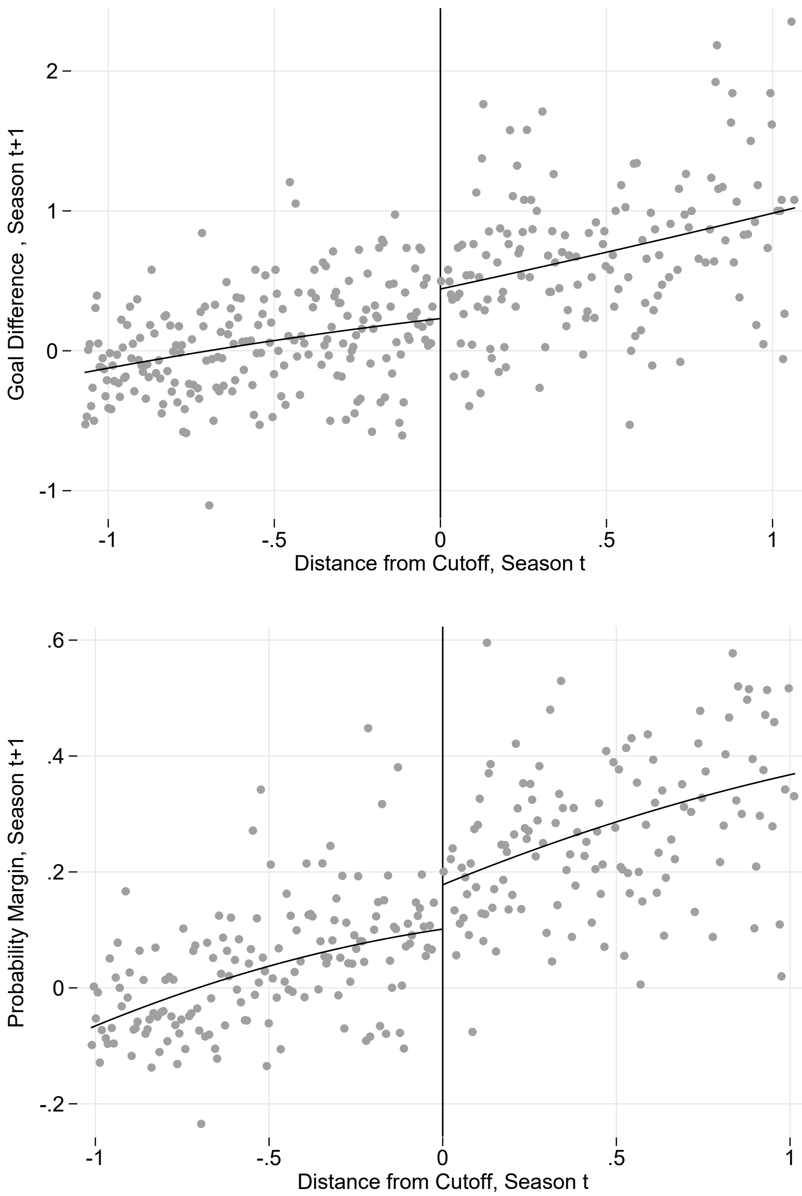
One way to quantify the magnitude of the discontinuity estimates is to consider how they scale relative to the national leagues champions. The discontinuity estimates of 0.3 in the goal difference corresponds to about 27% of the average goal difference that national league champions in season t achieve in season $t+1$. For the probability margin of winning, 10 percentage points improvement corresponds to approximately 25% of the average probability margin that national league winners achieve in season $t+1$.

A second approach to quantify the economic magnitude of the discontinuity estimates is to consider how causal effects may have affected the rankings of the teams close to the cutoff. In our sample, 5th ranked teams in season t have 0.24 goal difference per game in season $t+1$, whereas 4th and 3rd ranked teams have 0.37 and 0.68 goal differences. An increase of 0.3 goal difference per game would make a 5th ranked team to perform better than a 4th ranked team, but not better than a 3rd ranked team. Therefore, our coefficient estimate is economically significant as it alters the rankings of the teams meaningfully but it is not drastic as it push the 5th team (the team barely lost a UCL spot) by just 1 rank. For probability margin of winning, 5th, 4th, and 3rd ranked teams in season t have on average 0.11, 0.16, 0.26 probability margins in season $t+1$. Hence, a 10 percentage points increase in the probability margin of a 5th ranked team would potentially make the team perform better than the 4th ranked team in the following season. Thus, the economic magnitudes of our discontinuity estimates large.

6.2 Investigating Causal Channels

Table 3 underlies our estimates of the aggregate effect of participation in the UCL. As with any RDD study, these findings require further investigations about causal channels through which participation in the UCL affect performance. There are at least two reasons to think that elite sport tournaments improve the performance of the participant teams.

Figure 7: Discontinuity in Outcome Variables



Notes: Team's goal difference (top) and probability margin of winning (bottom) in season $t+1$, by distance from the cutoff in season t. Vertical lines indicate the cutoff, and dots indicate local averages. The solid lines are predicted values from quadratic polynomial on either sides of the cutoff.

Table 3: Discontinuity Estimates in Outcome Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Robust Bias-corrected				Conventional Method			
Dep. variable	GD(t+1)	GD(t+1)	PM(t+1)	PM(t+1)	GD(t+1)	GD(t+1)	PM(t+1)	PM(t+1)
Estimate	0.290**	0.314**	0.100***	0.116***	0.292***	0.312**	0.103***	0.111***
Std. Error	0.123	0.150	0.035	0.041	0.107	0.136	0.030	0.036
Bandwidth	0.760	1.075	0.611	1.015	0.760	1.075	0.611	1.015
Polynomial	1	2	1	2	1	2	1	2
Eff. Sample Size	20,281	29,231	16,544	27,827	20,281	29,231	16,544	27,827

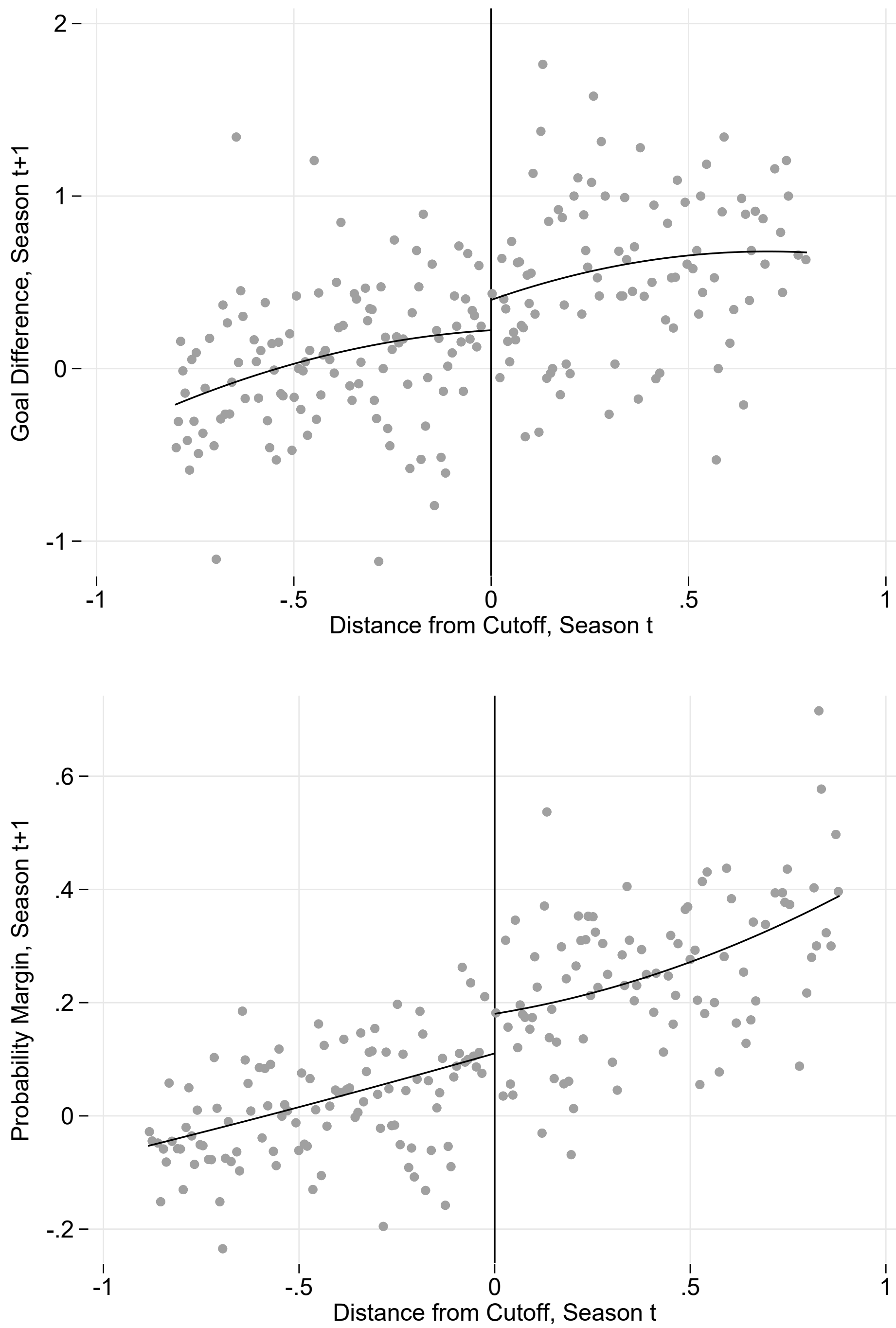
Notes: Estimates are based on linear and quadratic polynomial on each side of the cutoff, within a MSE-optimal bandwidth and triangular kernel. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

First, participants might learn from their peers or be motivated to practice harder, a mechanism referred to as "peer effect". Second, monetary benefits from the UCL might enable participating clubss to hire better players and managers, a mechanism referred to as "composition channel". While our identification strategy does not allow us to directly measure the peer effect, we can provide credible evidence that composition channel doesn't account for the improved performance of the UCL participant teams.

An important concern is that clubs that participate in the UCL have access to financial resources that may use to strengthen their team composition. In this section, we use the same regression discontinuity idea to investigate the team composition channel. In particular, we focus on two sub-channels through which participation in the UCL might work: (a) managerial changes, that is teams on two sides of the threshold may decide to change their managers, which might affect their performance; and (b) Player transfers, that is teams that qualify to play in the UCL may sign better quality players.

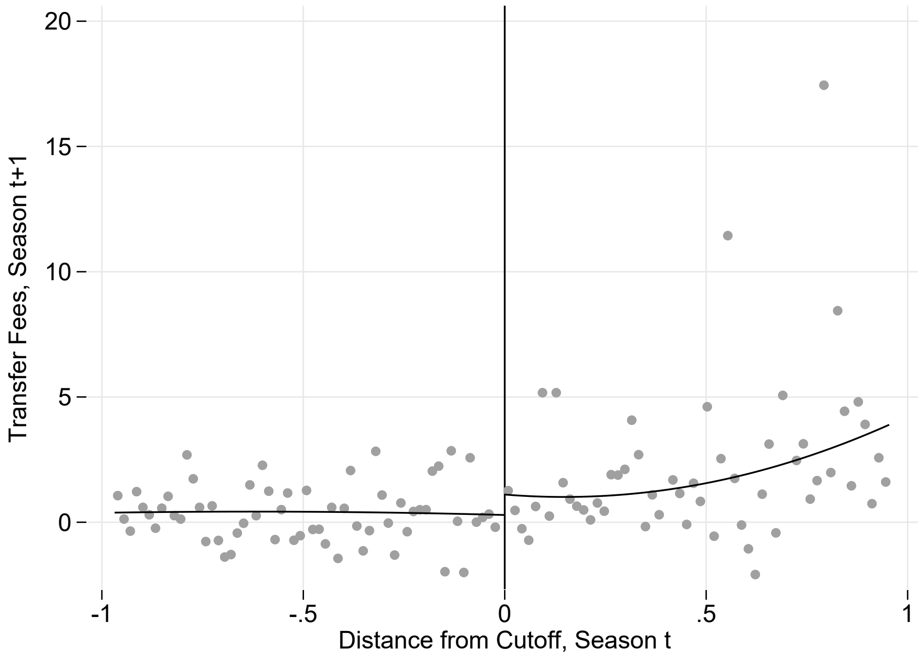
To investigate the effect of managers on team performance, we restrict the sample to the teams that had no managerial changes at the end of the season t (before the start of season $t+1$ and the UCL). If the improved performance is due to managerial changes, we should expect to find no discontinuity in team's performance in the sample

Figure 8: Discontinuity in Outcome Variables with No Managerial Changes



Notes: Team's goal difference (top) and probability margin of winning (bottom) in season $t+1$, by distance from the cutoff in season t. The sample is restricted to no managerial changes at the end of season t. Vertical lines indicate the cutoff, and dots indicate local averages. The solid lines are predicted values from quadratic polynomial on either sides of the cutoff.

Figure 9: Discontinuity in Transfer Fees



Notes: Transfer fees in season $t+1$, by distance from the cutoff in season t. Vertical lines indicate the cutoff, and dots indicate local averages. The solid lines are predicted values from quadratic polynomial on either sides of the cutoff.

with no managerial changes. Table 4, columns (1)-(2), reports the estimated total effect of participation in the UCL on the team's performance next season. Specifically, column (1) shows that the estimated discontinuity for the goal difference is 0.34 goals per game, about the same size as the full sample (see Table 3, column (1)-(2)). Similarly, column (2) shows that the estimated discontinuity for the probability margin of winning is 0.10 per game, almost identical to the full sample. These results suggest that managerial changes are unlikely to explain the improved performance reported in Table 3. A visual representation of these estimates is in Figure 8, which shows a jump in the goal difference and the probability margin of winning at the threshold. Therefore, we rule out

To rule out transfers as the other sub-channel, we look at the transfer fees on the two sides of the cutoff. Table 4, column (3) shows the estimated discontinuity for the

Table 4: Discontinuity Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Robust Bias-corrected			Conventional Method		
Dep. variable	GD(t+1)	PM(t+1)	Fees(t+1)	GD(t+1)	PM(t+1)	Fees(t+1)
Estimate	0.340*	0.100**	1.62	0.302*	0.100**	1.222
Std. Error	0.200	0.050	1.148	0.181	0.044	1.013
Bandwidth	0.805	0.889	0.969	0.805	0.889	0.969
Polynomial	2	2	2	2	2	2
Eff. Sample Size	15,809	17,711	7,100	15,809	17,711	7,100

Notes: Estimates are based on a quadratic polynomial on each side of the cutoff, within a MSE-optimal bandwidth and triangular kernel. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

transfer fees. Compare to the teams that do not participate in the UCL, teams that play in the UCL spend £1.62 million (in 2015 pounds) more on each transfer.[22] However, these estimates are not statistically significant at the 10% level. Figures 9 illustrates the same point by plotting transfer fees against the running variable. The figure confirms that the difference at the threshold is small and statistically insignificant, so it is not the case that the teams that narrowly qualify sign better quality players.

Overall, these results suggest that the causal impact of the UCL participation can be interpreted in terms of peer effects, rather than, for example, player transfers or managerial changes.

7 Conclusion

This paper analyzes whether and how participation in an elite sport tournament affects participating teams' performance. We ask whether competing against the best of other leagues improves performance of teams in their domestic leagues and if so how. To answer

[22]Our estimates (not reported) for total transfer fees per season is about £11 million in 2015 pounds, and statistically insignificant at 10% level.

our research questions, we compile a novel dataset of match level betting odds and goals scored. Using the betting odds, we construct an ex-ante measure of team performance at the match level, probability margin of winning: the extra probability betting market assigns for a team's win over its opponent's winning probability. Using goals scored information, we construct an ex-post measure of team performance at the match level, goal difference: the number of goals a team scored minus the number of goals it conceded. We link match level team performance measures with information on total points of teams at the end of season, eligibility and participation in the UCL.

We show causal effects of participation in the UCL on teams' performance in their domestic leagues with a fuzzy regression discontinuity design that exploits eligibility cutoffs. We identify a large and significant increase in the subsequent performance of the UCL participants. More specifically, teams that played in the UCL score about 0.3 goals per match more than teams that missed a UCL spot. Similarly, market assigns 10 percentage points more chance for the UCL teams to win a game than the teams not played in the UCL in that season.

The results we observe could be due to i) peer effects: teams getting better by competing against the best teams of Europe, ii) team composition changes: financial rewards of the UCL enabling teams to sign better players and to hire better managers. By ruling out team composition changes as a mechanism deriving these results, our results suggests the importance of peer effects. More specifically, transfer spending of UCL participants is not statistically higher than the non-participant teams close to the cutoff. Moreover, we still find positive and statistically significant effect even when we restrict our sample to clubs which did not change its manager over the summer.

Competing against the best requires every player in the team to be physically fit, the team to be well organized on the pitch, every player to be focused to the game on and off the pitch. Physical fitness and team organization on the pitch are developed through training sessions over the summer and throughout the season. Therefore, our hypothesis is that as teams expect tough competition in Europe, they take their training sessions more seriously and build up their physical and mental fitness and learn the team tactics. Enhanced physical and mental fitness, and adoption of team tactics helps team not just in the UCL but in their domestic leagues as well. Therefore, UCL participants achieve better outcomes in their domestic competitions than UCL non-participants.

References

Abdulkadiroğlu, A., Angrist, J., and Pathak, P., 2014. The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica*, 82(1):137–196.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., and Sterzi, V. Dancing with the stars: Innovation through interactions. Technical report, National Bureau of Economic Research, 2018.

Angrist, J. D., 2014. The perils of peer effects. *Labour Economics*, 30:98–108.

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

Bandiera, O., Barankay, I., and Rasul, I., 2010. Social incentives in the workplace. *The review of economic studies*, 77(2):417–458.

Buchdahl, J., 2016. *Squares & Sharps, Suckers & Sharks: The Science, Psychology & Philosophy of Gambling*, volume 16. Oldcastle Books.

Bürgi, C. and Sinclair, T. M., 2017. A nonparametric approach to identifying a subset of forecasters that outperforms the simple average. *Empirical Economics*, 53(1):101–115.

Calonico, S., Cattaneo, M. D., and Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Calonico, S., Cattaneo, M. D., and Titiunik, R., 2015. Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Carrell, S. E., Fullerton, R. L., and West, J. E., 2009. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464.

Cattaneo, M. D., Jansson, M., and Ma, X., 2019. Simple local polynomial density estimators. *Journal of the American Statistical Association*, pages 1–7.

Cornelissen, T., Dustmann, C., and Schönberg, U., 2017. Peer effects in the workplace. *American Economic Review*, 107(2):425–56.

Ehrenberg, R. G. and Bognanno, M. L., 1990. Do tournaments have incentive effects? *Journal of political Economy*, 98(6):1307–1324.

Falk, A. and Ichino, A., 2006. Clean evidence on peer effects. *Journal of labor economics*, 24(1):39–57.

Gelman, A. and Imbens, G., 2019. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3): 447–456.

Gould, E. D. and Winter, E., 2009. Interactions between workers and the technology of production: Evidence from professional baseball. *The Review of Economics and Statistics*, 91(1):188–200.

Guryan, J., Kroft, K., and Notowidigdo, M. J., 2009. Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68.

Hahn, J., Todd, P., and Van der Klaauw, W., 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Imbens, G. W. and Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

Lee, D. S., 2008. Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697.

Mas, A. and Moretti, E., 2009. Peers at work. *American Economic Review*, 99(1):112–45.

McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Moulton, B. R., 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, pages 334–338.

Sacerdote, B., 2001. Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics*, 116(2):681–704.

Triplett, N., 1898. The dynamogenic factors in pacemaking and competition. *The American journal of psychology*, 9(4):507–533.

UEFA. Regulations of the uefa champions league 2018-21 cycle, a. https://www.uefa.com/MultimediaFiles/Download/Regulations.

UEFA. How clubs' 2019/20 uefa champions league revenue will be shared, b. https://www.uefa.com/uefachampionsleague/news.

Waldinger, F., 2012. Peer effects in science: Evidence from the dismissal of scientists in nazi germany. *The Review of Economic Studies*, 79(2):838–861.

Zimmerman, S. D., 2019. Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review*, 109(1):1–47.

A Entries for the UEFA Champions League

In what follows we provide a brief overview of the national leagues from 2000 to 2019.

England: English Premier League (EPL) is the top level of the English football league, and contested by 20 clubs (380 matches per-season). The three lowest placed teams in the Premier League are relegated to the Championship, and the top two teams from the Championship promoted to the Premier League, with an additional team promoted after a series of play-offs involving the third, fourth, fifth and sixth placed clubs. During our sample period, the top 3-4 teams in EPL qualify for the UEFA Champions League (UCL).

Spain: La Liga is the top level of the Spanish football league, and is contested by 20 teams, with the three lowest-placed teams at the end of each season relegated to the Segunda División and replaced by the top three teams in that division. The top four teams in La Liga qualify for the UCL.

Italy: Serie A is the top level of the Italian football league, and is contested by 20 teams, with the three lowest-placed teams at the end of each season relegated to the Serie B and replaced by the top three teams in that division. During our sample period, the top 3-4 teams in Serie A qualify for the UCL.

Germany: Bundesliga is Germany's primary football competition and is contested by 18 teams, with the three lowest-placed teams at the end of each season relegated to the 2. Bundesliga and replaced by the top three teams in that division. During our sample period, the top 3-4 teams qualified for the UCL.

France: Ligue 1 is France's top division football competition and is contested by 20 teams,[23] with the three lowest-ranked teams at the end of each season relegated to the Ligue 2, and replaced by the top three teams in that division. During our sample period, the top 3 teams qualified for the UCL.

[23]Except for 2000-01 season where 18 teams were present.

Table 5: Entries for the UEFA Champions League Competition

	Football League				
Season	EPL	La Liga	Serie A	Bundesliga	League 1
2001-2002	20 (4,2,2)	20 (4,3,1)	20 (4,2,2)	18 (3,2,1)	18 (3,2,1)
2002-2003	20 (4,2,2)	20 (4,2,2)	20 (4,3,1)	18 (3,2,1)	20 (3,2,1)
2003-2004	20 (4,2,2)	20 (4,2,2)	20 (4,2,2)	18 (3,2,1)	20 (3,2,1)
2004-2005	20 (4,2,2)	20 (4,2,2)	20 (4,2,2)	18 (3,2,1)	20 (3,2,1)
2005-2006	20 (4,2,2)	20 (4,2,2)	20 (4,2,2)	18 (3,2,1)	20 (3,2,1)
2006-2007	20 (4,2,2)	20 (4,2,2)	20 (4,3,1)	18 (3,2,1)	20 (3,2,1)
2007-2008	20 (4,2,2)	20 (4,2,2)	20 (4,2,2)	18 (3,2,1)	20 (3,2,1)
2008-2009	20 (4,3,1)	20 (4,3,1)	20 (4,3,1)	18 (3,2,1)	20 (3,2,1)
2009-2010	20 (4,3,1)	20 (4,3,1)	20 (4,3,1)	18 (3,2,1)	20 (3,2,1)
2010-2011	20 (4,3,1)	20 (4,3,1)	20 (4,3,1)	18 (3,2,1)	20 (3,2,1)
2011-2012	20 (4,4,0)	20 (4,3,1)	20 (3,2,1)	18 (4,3,1)	20 (3,2,1)
2012-2013	20 (4,3,1)	20 (4,3,1)	20 (3,2,1)	18 (4,3,1)	20 (3,2,1)
2013-2014	20 (4,3,1)	20 (4,3,1)	20 (3,2,1)	18 (4,3,1)	20 (3,2,1)
2014-2015	20 (4,3,1)	20 (4,3,1)	20 (3,2,1)	18 (4,3,1)	20 (3,2,1)
2015-2016	20 (4,3,1)	20 (4,3,1)	20 (3,2,1)	18 (4,3,1)	20 (3,2,1)
2016-2017	20 (4,3,1)	20 (4,3,1)	20 (3,2,1)	18 (4,3,1)	20 (3,2,1)
2017-2018	20 (4,4,0)	20 (4,4,0)	20 (4,4,0)	18 (4,4,0)	20 (3,3,0)
2018-2019	20 (4,4,0)	20 (4,4,0)	20 (4,4,0)	18 (4,4,0)	20 (3,3,0)

Notes: Each entry indicates the number of teams in the League in a season from a specific league. The numbers in the paranthesis indicate the number of teams that are eligible, the number of eligible teams that directly participate in the UCL group stage, and the number of teams that play in the play-off rounds to qualify for the group stage, respectively. *Source: Wikipedia.*

B Robustness Checks

Table 6: Discontinuity Estimates in Predetermined Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Robust Bias-corrected				Conventional Method			
Dep. variable	GD(t-1)	GD(t-1)	PM(t-1)	PM(t-1)	GD(t-1)	GD(t-1)	PM(t-1)	PM(t-1)
Estimate	0.150	0.132	0.063	0.065	0.258	0.140	0.092	0.060
Std. Error	0.159	0.195	0.052	0.064	0.139	0.170	0.044	0.055
Bandwidth	∞	∞	∞	∞	∞	∞	∞	∞
Polynomial	2	3	2	3	2	3	2	3
Eff. Sample Size	43,217	43,217	43,217	43,217	43,217	43,217	43,217	43,217

Notes: Estimates are based on a quadratic and cubic polynomial on each side of the cutoff. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 7: Discontinuity Estimates in Outcome Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Robust Bias-corrected				Conventional Method			
Dep. variable	GD(t+1)	GD(t+1)	PM(t+1)	PM(t+1)	GD(t+1)	GD(t+1)	PM(t+1)	PM(t+1)
Estimate	0.389***	0.292*	0.121***	0.112***	0.24**	0.40***	0.089***	0.124***
Std. Error	0.118	0.149	0.032	0.040	0.098	0.129	0.026	0.034
Bandwidth	∞	∞	∞	∞	∞	∞	∞	∞
Polynomial	2	3	2	3	2	3	2	3
Eff. Sample Size	54,276	54,276	54,276	54,276	54,276	54,276	54,276	54,276

Notes: Estimates are based on a quadratic and cubic polynomial on each side of the cutoff. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 8: Discontinuity Estimates in Outcome Variables with No Managerial Changes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Robust Bias-corrected				Conventional Method			
Dep. variable	GD(t+1)	GD(t+1)	PM(t+1)	PM(t+1)	GD(t+1)	GD(t+1)	PM(t+1)	PM(t+1)
Estimate	0.357**	0.142	0.115***	0.080**	0.210*	0.371**	0.086***	0.118***
Std. Error	0.140	0.177	0.036	0.046	0.115	0.154	0.030	0.034
Bandwidth	∞	∞	∞	∞	∞	∞	∞	∞
Polynomial	2	3	2	3	2	3	2	3
Eff. Sample Size	38,559	38,559	38,559	38,559	38,559	38,559	38,559	38,559

Notes: Estimates are based on a quadratic and cubic polynomial on each side of the cutoff. The sample includes the observations with no managerial changes at the end of season t. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 9: Discontinuity Estimates in Transfer Fees

	(1)	(2)	(4)	(5)
	Robust Bias-corrected		Conventional Method	
Dep. variable	Fees(t+1)	Fees(t+1)	Fees(t+1)	Fees(t+1)
Estimate	0.768	1.32	0.444	0.799
Std. Error	0.870	1.059	0.715	0.094
Bandwidth	∞	∞	∞	∞
Polynomial	2	3	2	3
Eff. Sample Size	13,577	13,577	13,577	13,577

Notes: Estimates are based on a quadratic and cubic polynomial on each side of the cutoff. All specifications include a season and league fixed effects. Estimated standard errors are two-way clustered at the team-season levels. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.