Favoritism and Social Pressure Revisited: Bowing to Power, Not the Crowds

Gábor Békés*[1,2,3], Endre Borza[2], and Márton Fleck[1]

[1] *Central European University, Vienna, Austria*
[2] *Centre for Economic and Regional Studies, Budapest, Hungary*
[3] *C.E.P.R., London, United Kingdom*

Draft, v1.24– 21 September 2022. See latest version HERE.

Abstract

We contribute to the literature by showing that telling apart sources of social pressure is important. We revisit Garicano, Palacios-Huerta, and Prendergast (2005) who analyzed the decisions of association football (soccer) referees and found evidence of favoritism: referees helping the home team in crucially important situations. They argued that this bias is driven by social pressure stemming from home team supporting crowds in the stadium: referees show conformity to the home team supporters. We argue that referees may be affected by multiple sources of social pressure: in particular supporting crowd in the stadium as well as the host team organization. Relying on exogenous variation in crowd size, we find no effect from crowds on referee behavior. Instead, we show that the host team organization may have an influence on referees as mostly successful teams benefit from the referees' favoritism bias. Nevertheless, social pressure remains the driver of favoritism, as we find no evidence of corruption – bias in favor of successful teams is uncorrelated with referee career.

Keywords: social pressure, favoritism, persuasion, corruption, sport

JEL-codes: D71, Z20, C21

1 Introduction

In a situation of competition between organizations, there is often a role for an agent (such as a judge or referee) to make a decision benefiting one party. Favoritism is the practice when a decision maker gives preferential treatment to one party at the expense of another, which is not warranted by the rightful determinants. Favoritism is present in a wide variety of situations ranging from promotion decisions in organizations (Prendergast and Topel, 1996) to allocation of regional development funds (Hodler and Raschky, 2014).

*Corresponding author. Central European University, Quellenstrasse 51, Vienna, Austria, bekesg@ceu.edu. We thank Mats Koster, Miklós Koren, Ádám Szeidl, László Mátyás, and Thomas Peeters and seminar participants at CEU for useful comments and suggestions.

Favoritism may stem from explicit corruption (bribery) while one source of favoritism is social pressure, the exertion of influence by a person or a group. Social pressure includes conformity and persuasion (excess impact of the beliefs of others). Conformity is about the agents' desire to conform to some expectations, to seek approval or adhere to a social image of themselves (Bursztyn and Jensen, 2017). Persuasion occurs when individuals or groups try to influence an agent's decision to their interest via persuasive behavior such as communication (DellaVigna and Gentzkow, 2010) looking for both a cognitive and emotional response from the receiver (DeMarzo, Vayanos, and Zwiebel, 2003; Schwartzstein and Sunderam, 2021). When social pressure leads to favoritism, policy actions should raise awareness and aim to reduce bias by affecting either party (the sender or the receiver). This is easy when there is a single channel of social pressure. However, individuals may care about perception by more than one reference group (Bursztyn and Jensen, 2017). For instance students may care about perception both by their peer group as well as by potential employers (Austen-Smith and Fryer, 2005). Different parties may aim at persuading consumers at once, too.

We contribute to the literature by showing that telling apart sources of social pressure is important. We revisit Garicano et al. (2005) who analyzed the decisions of association football (soccer) referees and found evidence of favoritism: referees helping the home team in crucially important situations.[1] They argued that this bias is driven by social pressure stemming from home team supporting crowds in the stadium: referees show conformity to the home team supporters and thus, internalize their preference. We will argue that referees may be affected by multiple sources of social pressure: in particular supporting crowd in the stadium as well as the host team organization. Relying on exogenous variation in crowd size, we will find no effect from crowds on referee behavior. Instead, we will show that the host team organization may have an influence on referees as mostly successful teams benefit from the referees' favoritism bias. Nevertheless, social pressure remains the driver of favoritism, as we find no evidence of corruption – bias in favor of successful teams is uncorrelated with referee career.

To be able to separate sources we need appropriate data and an identification strategy that can help the differentiation. To that end, first, we compiled a new and comprehensive dataset on European football games covering match events, referee decisions, and referee career paths. Our high granularity dataset covers 10 seasons (2011/12 to 2020/21) of the 5 most prestigious men's football leagues: the English *Premier League*, the Spanish *La Liga*, the Italian *Serie A*, the German *Bundesliga 1*, and the French *Ligue 1* first divisions. We work with event-by-event data that records each action (such as passes, disciplinary action, penalties, injuries) along with a timestamp of that event.[2]

This paper confirms the existence of the favoritism result by Garicano et al. (2005): referees use their discretionary power to systematically award more stoppage time to the home team

[1] There is a growing literature using sports to learn about behavior. For instance, expectation of financial rewards was shown to lead to match rigging in Japanese Sumo wrestlers (Duggan and Levitt, 2002). Gauriot and Page (2019) also use football data to talk about quality perception biased by luck, while Parsons, Sulaeman, Yates, and Hamermesh (2011) look at discrimination in US baseball.

[2] Note that the bias could be captured by looking at disciplinary action such as awarding penalties or yellow cards. While these are easier to observe, identification is problematic, as these decisions will be intertwined with the actions of the players making it difficult to empirically disentangle referee and player behavior, see e.g. Carmichael and Thomas (2005); Dawson and Dobson (2010), and for a review, Dohmen and Sauermann (2016).

when it stands to benefit.[3] Comparing games with a one goal difference after regular time, we find that referees on average add 13 seconds more stoppage time when the home team is losing compared to when the home team is winning.[4]

Next we turn to understanding the sources of social pressure. First, we examine the role of the supporting crowds. Garicano et al. (2005) argued that referee's home bias is driven by social pressure stemming from home team supporting crowds at large in the stadium: referees show conformity and internalize the preferences of the home team supporters. Conformity happens as referees pursue to "satisfy the crowds": gain their approval and internalize their preferences. This mechanism is identified via variation in crowd size and composition: larger attendance generate greater bias, larger than usual capacity – a posited sign of higher share of visiting crowds – yields lower bias. These findings are extended by Dohmen (2008) who suggested that distance to crowds – in stadiums with running tracks – moderated the influence of crowds.

We find, however, that the bias is not driven by crowds. The global Covid pandemic in 2020 and 2021 led to sudden and unexpected stadium closures, and games were played in empty stadiums. Instead of relying on observational variation in relative size of home supporters, we could use stadium closures as a natural experiment. Using this exogenous variation in home support by local crowds, we found that the bias remained unchanged even without supporters in the stadium.

The home team's organization is the second source we look at. The organization hosts referees, who might experience social pressure from them. Persuasion could occur with the host team having the chance to repeat a positive message, and hospitality may play an important role as well. On the day of the match, referees are surrounded by members of the host organization and often spend the whole day at the premises. Before, during, and after the match, the home team staff accounts for the safety and well-being of the referee. Gifts from the home team are not uncommon either.[5] At the same time, referees in close proximity to stars and powerful managers may also want to conform to teams.

Social pressure is likely to stem from both teams, but its relative size will vary with influence. In football, the size of influence is related to sporting success (such as ranking after a season). For the referees, the desire to conform to successful teams and their star players will be stronger. These teams will also be able to mobilize more resources (better facilities and personnel) for the purpose of persuasion.[6] To detect social pressure related to the host team organization, we will compare teams of different ranking: more successful teams are expected to experience a larger home team bias.

Looking at the difference in stoppage time when it benefits the home team compared to when

[3]The discretionary power empowers referees to set the stoppage time at the end of each half of the game to compensate for the time lost for a variety of events such as injuries.

[4]This difference corresponds to 5% longer stoppage time, or an average of one additional point per season and team. This is somewhat smaller than the 20 second bias found by Dohmen (2008) in the German Bundesliga for the 1992/93 to 2003/04 period, and substantially smaller than the 110 seconds Garicano et al. (2005) found for the two Spanish seasons of 1994/95 and 1998/99.

[5]For anecdotal evidence, two examples are `https://www.theguardian.com/football/2015/mar/28/referees-football-match-day-routine-sport`, `https://en.as.com/en/2017/02/27/soccer/1488229755_283818.html`.

[6]Sporting success and team wealth are strongly correlated in elite sport like football, and there is low churning at the top.

benefits the away team, we indeed find that referees add more than twice as long stoppage time for the benefit of the most successful third of teams compared to the rest.[7] The favoritism bias is the largest when a top ranked team plays a minnow: the lowest ranked teams will enjoy zero home bias when playing against top teams.

In our exercise, we aimed at showing an example when social pressure may come from different sources. We found that the traditionally assumed source (crowds) cannot explain favoritism. How can we reconcile our results with earlier ones? We suspect that earlier results on crowds may have been confounded by characteristics of the host team organization that remain in play even behind closed doors: more successful teams have larger stadiums, and are less likely to have running tracks (no top team in Germany has tracks, while 9.8% of others have tracks). A large and fine dataset and exogenous variation was necessary to distinguish these two sources of social pressure.

Importantly in this setup, social pressure may have an aggregate consequence. If the top ranked teams get additional help from the referees, social pressure will contribute to maintaining the ranking thereby making it more difficult for smaller teams to catch up.

Finally, we examine whether there could be a more standard explanation to favoritism that could, at the same time, act as a key possible confounder: corruption. While direct bribery is extremely unlikely,[8] referees may expect professional rewards in the future when helping the most successful teams.[9] To refute this alternative hypothesis, we examined referees' careers in the prestigious pan-European competitions where only the most successful teams play. If favoritism was driven by the expectation to work more in these leagues, referees with higher bias would be picked more – in collusion with participating teams. We find no such evidence.

This paper contributes to understanding the importance that social pressure may be related to several actors at once, and their identification will have different policy outcomes. Consider the case of possibly biased arbitration judgements in investor-state dispute settlements (Behn, Berge, and Langford, 2018). Here there may be a variety of social pressures: persuasion by and conformity to wealthy countries as well crowd pressure via social media or protests. In the case of investigating bias in online reviews (Vollaard and van Ours, 2022), reviewers may be subject to social pressure when favoring big brands popular with the largest communities as well as be the target of persuasion by companies sending gifts and offering marketing events. In both these scenarios our results suggest that an analysis must try tackle different sources rather then assuming either.

As we are interested in broader social settings, we acknowledge that the football setting is particular. However, all games are televised, and detailed data (like the one that we use) is shared in real time and such transparency should minimize any biased behavior. Thus, in other social and economic settings with a lack of public attention, one may expect a higher bias.

In what follows, we first describe the dataset we used and key empirical methods we applied

[7] This is in line with a related finding: using an expert panel, Erikstad and Johansen (2020) showed that the top two teams in Norway are more likely to get a penalty awarded.

[8] Although not unheard of, see the Italian match fixing case `https://en.wikipedia.org/wiki/Calciopoli`

[9] For instance, the most successful teams are able to pressure the UEFA to adapt its competition formats in their favor or avoid penalties. For some discussion, see Appendix Section 5.3.

in Section 2. Then we discuss our results step by step in Section 3, before a brief summary in Section 4.

2 Data and Empirical Strategy

In this section, we first present our dataset, which has been compiled from several sources.[10] Second, we describe our core empirical model and the variables we used.

2.1 Data Sources

Our main dataset covers the universe of men's football matches of the top five European leagues (England, France, Germany, Italy, Spain). In a season, each team plays every other team twice, once at home and once as visitor. Over the period of 10 seasons (from 2011/12 to 2020/21), we have $N = 18,118$ matches.[11]

Such dataset has multiple advantages. Multiple leagues allow filtering out possible country specific rules and customs and lead to high external validity, and a large coverage is also necessary to power our identification. Detailed and long time series information is necessary to examine referees careers.

The main dataset is an event-by-event level description of every game, collected from `whoscored.com`. Each event has a type and a timestamp (period, minute and second). Event types include: pass, ball recovery, throw-in, free-kick, yellow and red card, substitution, penalties, injury, shot, goal, or corner. In a typical game, an event happens once in 3.6 seconds: there are 1432 events recorded in the regular playtime of 90 minutes, and 90.1 events during stoppage time.

The second set of data is at the game level. It includes the venue of the game, attendance in the stadium, the result (goals by home and away teams), referee name, date, and time.

The third set of data concerns referees' experience in terms of the number of games they refereed in domestic and European competitions. The data has been collected from `soccerway.com`.

We used complementary data from a variety of sources: information on stadiums comes from `transfermarkt.com` as well as from Wikipedia pages of teams. Pre-match odds come from of the betting site `bet365.com`, financial power is based on revenues according to the Deloitte Football Money League, and squad values are based on valuations from `transfermarkt.com`.

2.2 Empirical Strategy

Our main outcome variable of interest, $Stoppage\_time$, is the length of playtime (measured in seconds) beyond the regular time (90:00 minutes). This second-half stoppage time compensates for wasted time in the second half of the game. Its length is calculated from the the event data based on the timestamp of the event that indicates the end of the game.

[10] Data and code will be made available.

[11] There are 20 teams in a league (18 in Germany), $10 \times 4 \times 20 \times 19 + 10 \times 1 \times 18 \times 17 = 18260$. Due to the Covid-19 pandemic, the season 2019/20 of the French Ligue 1 finished early, with only 279 out of 380 games played. Due to data coverage deficiencies, we lost 41 games.

We follow Garicano et al. (2005), and focus only on the cleanest comparison to study referees' favoritism bias: looking at matches where the goal difference at the end of the regular time (90:00) is exactly one goal. Our key dependent variable is an indicator variable, $Home\_lose$, as we compare the length of stoppage time in games when the home team is losing by one goal ($Home\_lose = 1$) with games where it is winning by one goal ($Home\_lose = 0$). In our estimation sample, we have $N = 7,021$ games (out of 18,118). For each game, events are aggregated at the first half, the second half, as well as the first and second stoppage time periods.

The average stoppage time is 260 seconds, ranging from 3 to 1065 seconds. The home team wins slightly more games than the away team (56% compared to 44%), consistent with the well-documented general home advantage in sport (Jamieson, 2010). For a broader review of descriptive statistics, see Table 5 in the Appendix. This difference between average stoppage time by the home team losing or winning, however, could be confounded by a variety of factors, such as injuries correlated with both stoppage time and the result, or differences in playing style of the home team. To partial these out, we estimate the following model with OLS:

$$Stoppage\_time_{h,a,s} = \beta Home\ lose_{h,a,s} + \gamma Controls_{h,a,s} + \eta_h + \epsilon_{h,a,s}, \tag{1}$$

where our unit of observation is a single game played between home team h, and away team a in season s. As each team hosts every other team once in a season, the h, a dyad uniquely identifies a game in any season s. We use a rich set of control variables ($Controls_{h,a,s}$) as described below. Standard errors are clustered at the home team level[12].

First, we approximate the justifiable length of stoppage time by counting the time during which the ball is likely to have been out of play. *Wasted time* is a sum of seconds between two consecutive events if the first event is a foul, a card, a ball picked up by the keeper, or a goal; or if the second event is a corner, a throw-in, or a substitution. In the second half, wasted time varies between 8.5 and 33.5 minutes, its average is 19 minutes (out of 45+4.5), equivalent to 62% effective playing time.

Second, in addition to wasted time, we also control for the number of events associated with long stops: yellow or red cards, substitutions, fouls, and goals in the second half. It may matter as referees may use heuristics such as the number of these key events to decide stoppage time.

Furthermore, these events are potentially confounding variables as they can be correlated with the goal difference as well, given that the playing style of the teams usually varies depending on winning or losing.

Third, a potential confounding effect may be that instead of favoring the home team, referees may simply let the attacks started during the end of the stoppage time finish. It is a confounder because in general, the losing team is likely to play more offensively during stoppage time (as they need to score a goal), and the away team is more likely to lose (due to the home advantage in general). To control for this possibility, we generate a variable by taking into account the passes of the losing team during stoppage time, and take the average distance of these passes from the team's own goal line. This variable (called *Losing offensiveness*) is measured in units of distance from the team's own goal line, on a scale from 0 to 100.

[12]Standard errors with the alternative home-away level clusters are slightly smaller

Fourth, during the examined 10-year period, the video assistant referee (VAR) technology was introduced, and its use may affect both the stoppage time setting and the activity of players. Thus, we added a league-season level variable (*VAR*) indicating whether the technology was in use.[13]

Fifth, for each match we also control for the round of the season. This variable runs from 1 (the first match in the season for each team) to 38 (the last match in the season for each team in case of a league with 20 teams). In later rounds there is more stake in the game that can affect both the referee's and team's behavior.

Finally, in line with earlier literature, we add home team fixed effects η_h to capture the footballing style and quality of the team. This allows within team comparisons of stoppage time conditional on end of regular time result. Our baseline specification does not include referee fixed effects, as we cannot rule out the possibility that the allocation of referees is not random, and may be a part of the mechanism through which bias works. As shown later, this decision has no practical consequence.

3 Results

This section presents our empirical findings. We confirm the existence of a referees' favoritism bias towards the home team, and show that this bias is not driven by the fans in the stadium. As a next step, we show that influential teams enjoy a larger favoritism bias from referees. Finally, we show that the home bias of referees is not associated with their career, thus we can rule out *quid pro quo* motives behind the bias.

3.1 Home-team Bias is Still There, But Not Because of the Crowd

In our dataset, 14.28 seconds is the raw difference between the additional stoppage time if the home team is losing by one goal and the additional time when it is winning by one goal. As shown in Table 1,once all control variables are added, this difference is marginally changed to 12.80. For more details about the distribution of stoppage time, and how controls affect the estimated coefficient, see Appendix Section 5.1.[14]

Due to the Covid-19 outbreak in Europe in Spring 2020, practically every football league was suspended as of the second weekend of March. France closed the 2019/2020 season early, while other leagues resumed around May-June, with matches played behind closed doors. Closed door games continued in 2020/21 season and in Spring 2021, partial opening meant that for 139 games overall stadiums were filled to an average capacity of 13%. In these two years, about 2/3 of games were behind fully or partially closed stadiums. (For details, see Appendix Table 4 in

[13] VAR has been in operation since season 2017/18 in Germany and Italy, since season 2018/19 in France and Spain, and since season 2019/20 in England. See more in the Appendix Section 5.1.

[14] Our results refer to the 2011-2021 period. Note that our estimated coefficient for Germany is 16 seconds, close to what was measured earlier by Dohmen (2008) but a magnitude smaller than the one in Garicano et al. (2005). In lack of available data from the period, it is difficult to make a direct comparison, but the difference in estimates is not driven by modelling choices: a replication of their core model (in their Table 2, column 4) offers a similar estimate to our favored specification (Appendix, Table 8). It is possible that in the nineties referee behavior in Spain had less oversight.

Table 1: Presence of home bias, no crowd effect

	(1)	(2)	(3)	(4)
Home lose	14.28***	12.80***	13.22***	13.17***
	(2.24)	(1.79)	(1.77)	(1.77)
Home lose × Covid			-1.35	
			(5.46)	
Home lose × Closed				-0.71
				(5.44)
Controls	No	Yes	Yes	Yes
League FE	No	Yes	Yes	Yes
Home team FE	No	Yes	Yes	Yes
R^2	0.01	0.41	0.41	0.41
Observations	7021	7021	7021	7021

Notes: Games with a single goal difference after regular time, N=7021. Dependent variable is stoppage time in seconds. Controls include time with ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, round of season, whether VAR was used, and the average distance of the passes of the losing team from opponent's goal line in the extra time. In columns 3 and 4, all controls are also interacted with the Covid or the Closed dummy. Standard errors clustered at home team level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the Appendix.) We created two indicator variables, $Closed = 1$ when attendance is zero, and $Covid = 1$ that also include very low attendance games in 2021. As partial opening only affects 6.6% of games, it will turn out to have very little impact.

To find out if the difference disappeared during closed games, we added an interaction term to 1 and estimated:

$$\begin{aligned} Stoppage\_time_{h,a,s} = \beta_1 Home\ lose_{h,a,s} + \beta_2 Closed_{h,a,s} + \beta_3 Home\ lose_{h,a,s} \times T_{h,a,s} \\ + \gamma Controls_{h,a,s} + \eta_{hometeam_h} + \epsilon_{h,a,s}, \end{aligned} \tag{2}$$

where our treatment indicator $Closed_{h,a,s}$ flags closed matches. An alternative is where instead, we flag all closed and partially open games ($Covid_{h,a,s}$). All control variables are also interacted with the *Closed* (or *Covid*) dummy to capture that without crowds, players may behave differently or referees may take a different amount of time to make decisions.

Columns 3 and 4 in Table 1 show that the bias is unchanged even in fully or mostly empty stadiums: the estimated interaction terms are very close to zero. Once again, this result is rather robust to changing the mix of control variables (including adding referee fixed effects), see Table 7 in Appendix. It is also stable across across leagues – see Figure 5 in Section 5.8 in the Appendix.

The exogenous variation in crowd presence allowed us to test the hypothesis that the favoritism bias from referees favoring the home team is the consequence of the social pressure exerted by the fans of the home team in the stadium. Our results confirmed bias in referee decisions, but can rule out that this is driven by the size of the crowd, as it unchanged even in the extreme case of closed stadiums. This is our first main result: differences in the crowd size is very unlikely to be the mechanism behind the social pressure, referees must be helping the home team for other reasons.

3.2 The Home-team Bias Driven by Influential Teams

If the favoritism bias from referees is not driven by crowds, there must be some other mechanism that leads to biased behavior.

In this section we investigate if the size of the home-team bias is correlated with team influence. Sporting success and financial clout allows teams to have influence: attract players, fans, investment, media interest and so on. Influence will also provide teams the capacity to exert social pressure in the form of persuasion of independent agents.

Financial and sporting success are strongly correlated: wealthier teams will have better players and will win more. In our baseline specification we use the league table ranking as it is a well-defined order of team success. In our data, wealth and quality measures are highly correlated with a correlation coefficient between 57% and 92%. In a robustness check, we use replicate results with other metrics of influence.

Ranking is defined as the end-of-season position of the team in the league table, the lower the better (1 is the winner, 18 or 20 is the last team). The final ranking of a team at the end of a year is close to its expected average ranking throughout the year, and it may be easily compared across leagues and seasons.

To uncover the relationship between stoppage time and ranking, we first estimate a model with only the football rule controls (such as time with the ball out of play, number of cards, substitutions), as described in Section 3.1. Then, we compute the difference between predicted and actual stoppage time. This deviation, $\hat{Bias}$ from equation 3, is the measure of the unexplained difference.

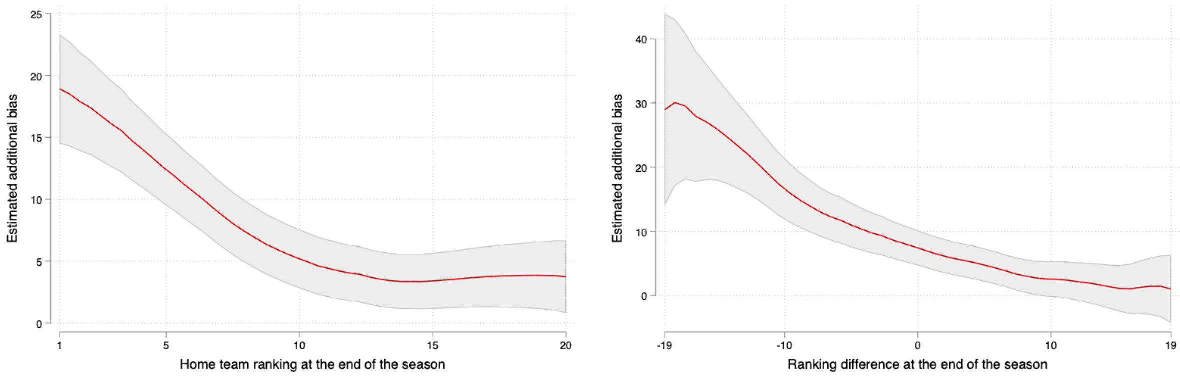
$$\hat{Bias} = Stoppage\_time_{h,a,s} - (\hat{\gamma} Controls_{h,a,s} + \hat{\eta}_{hometeam_h}) \tag{3}$$

In the second step, using local polynomial smoothing regressions, we plot this bias against heterogeneity by the ranking of the home team. In Figure 1a, we see a fairly strong pattern with the higher-ranked teams enjoying a greater home-team bias. In Figure 1b, we see a similar pattern with the difference plotted against bias: when the top team plays against the lowest ranked one, the gap is 30 seconds, but it goes down to 0 when a minnow plays at home against a top team.[15]

Both graphs suggest that influence in terms of ranking may be non-linear. This may stem from an important rule: the top six teams leads to playing in the rewarding European championships next year. Thus, we will also consider a binary variable, *Top6*, to measure influence being inside

[15] Alternatively, confounders may be partialled out of rank as well, only to give very similar graph.

Figure 1: Heterogeneity of home-team bias by team ranking



(a) Home team ranking (b) Difference between home and away team ranking

Notes: Local polynomial smoothing and a 95% confidence interval. Bias measured in seconds, rank between 1 (best) and 20 (worst); 1-18 for Germany. Predicted stoppage time is the residual from a regression of stoppage time on football rule controls (time with ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, round of season, whether VAR was used, losing offensiveness (as average distance of the passes of the losing team from opponent's goal line in the stoppage time) all interacted with *Closed* dummy.

or outside the top six teams per league.

We estimate this heterogeneity in two ways. First, we investigate difference along team influence (here: ranking) with β_3 in Equation 4 measuring the heterogeneity in home-team bias. The $Home\_ranking_{h,s}$ variable may be estimated in a linear form or with a binary Top6 variable.

$$\begin{aligned} Stoppage\_time_{h,a,s} = {} & \beta_1 Home\_lose_{h,a,s} \\ & + \beta_2 Home\_ranking_{h,s} + \beta_3 Home\_lose_{h,a,s} \times Home\_ranking_{h,s} \\ & + \gamma Controls_{h,a,s} + \eta_{hometeam_h} + \epsilon_{h,a,s}, \end{aligned} \tag{4}$$

Second, an alternative model uses the difference between home and away team ranking, with β_3 measuring the heterogeneity in home-team bias in terms of the difference between ranking. The $Home\_ranking\_difference_{h,s}$ variable may be estimated in a linear form or with a binary variable, where high difference is defined as a gap greater than -10. In both cases note that home team fixed effects soak any time invariant asp

$$\begin{aligned} Stoppage\_time_{h,a,s} = {} & \beta_1 Home\_lose_{h,a,s} \\ & + \beta_2 Home\_ranking\_difference_{h,a,s} \\ & + \beta_3 Home\_lose_{h,a,s} \times Home\_ranking\_difference_{h,a,s} \\ & + \gamma Controls_{h,a,s} + \eta_{hometeam_h} + \epsilon_{h,a,s}, \end{aligned} \tag{5}$$

The results of both models are presented in Table 2. In the simplest setup with the Top6 indicator in column (1) of Table 2, we see that influential teams enjoy a home-team bias that is more than twice the one for the rest (10.23 vs 22.98 seconds). Looking at the home team's ranking in a linear way in Column (2), we see $20.48 - 1 \times 0.67 = 19.80$ seconds for the top team, reduced by 0.67 seconds per rank, shrinking to 7.10 seconds for the lowest ranked one.

Next, we add the difference between teams in terms of ranking from the perspective of the home team. The difference thus ranges between 19 (when the lowest ranked team plays at home against the best one) and -19 (when the top team hosts the bottom team).

In the binary setting (Column 3), we once again find that when there is a sizeable rank gap in favor of the home team, the bias is twice the size compared to a case with no difference (12.05 vs 12.05+15.65=27.70 seconds). In the linear model (Column 4), for equally ranked teams the bias is 13.94 seconds and the slope is -0.84. So the lowest ranked team actually had a -2-second (ie. negative) bias when playing against the top team, while the top team enjoyed, a 29.9-second benefit against the lowest one.

Table 2: Regressions indicating home-team bias heterogeneity

	(1)	(2)	(3)	(4)
Home lose	10.23***	20.48***	12.05***	13.94***
	(1.84)	(4.41)	(1.76)	(1.86)
Home lose × Home Top 6	12.75***			
	(3.74)			
Home lose × Home rank		-0.67**		
		(0.33)		
Home lose × Home-away rank diff ≤ -10			15.65**	
			(6.46)	
Home lose × Home-away rank diff				-0.84***
				(0.21)
Controls	Yes	Yes	Yes	Yes
League FE	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes
R^2	0.41	0.41	0.41	0.42
Observations	7021	7021	7021	7021

Note: Games with a single goal difference after regular time, N=7021. Dependent variable is stoppage time in seconds. Controls include time with ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, round of season, whether VAR was used, and the average distance of the passes of the losing team from opponent's goal line in the stoppage time. Control variables are also interacted with the *closed* dummy. Standard errors clustered at home team level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

These results suggest that referees have a home-team bias that is substantially higher for more influential (top-ranked) teams, especially when they play against less influential (low-ranked) ones.

As noted earlier, influence captures aspects of financial and sporting success, and ranking is not the only way to measure it. We considered three alternatives. First, we used pre-game odds values instead of ranking to better capture the perception of influence. Second, in relation to wealth, we defined top teams as the 20 financially richest teams (defined by Deloitte Football Money League) across these five leagues for the period in question. Third, instead of rank difference, we considered difference in team squad values. We found our results qualitatively unchanged when using these alternative definitions. For details, see Section 5.9 with Table 12 in the Appendix.

To summarize, we find that influential (top-ranked) teams benefit substantially more from referee decisions. When playing at home, the top teams seem to create an environment that makes referees more inclined to help. It is the influence of the host team organization and not the crowd size that affects referee behavior.

Our findings may explain earlier evidence in the literature suggesting crowd and stadium characteristics affecting referee decisions. Team influence is correlated with many observable characteristics: stadium attendance and capacity or the type of stadium teams have. In our data, 75% of variation in attendance rate, an earlier measure of the crowd effect (Garicano et al., 2005) is explained by teams, league, and season (see Table 9 in the Appendix). Similarly for running tracks (Dohmen and Sauermann, 2016), the top 6 teams in Germany have no tracks in their stadium but some of the rest (like Hertha Berlin, or Nuernberg) do. Hence, crowd influence lose its role as explanation once confounders are taken into consideration.

3.3 Is There Corruption? Referee Career in Europe

Finally, we investigate if career concerns may motivate referees to help influential teams. The specific question we address here is whether more biased referees are more likely to get work in the UEFA Champions League and in the UEFA Europa League games. Refereeing in these European games is a pinnacle of a referee's career, and the number games they worked at is a key career success metric. Influential teams – especially those with regular presence in these competitions also having positions at UEFA boards – can (albeit rather informally) block unwanted referees and possibly promote preferred ones.[16]

To analyze careers, we aggregated game information into an unbalanced referee-season panel ($N = 1,148$) of referees ($N = 233$) and seasons ($N = 10$, $s = 2011-12 \ldots 2020-21$). On average we observe a referee for 4.9 seasons in their national leagues working on 70 games, 33 of which have a one-goal difference after 90 minutes.

We keep only referee-season pairs where we observe at least one 0:1 and 1:0 result after regular time and we are left with 179 referees and $N = 711$ observations. This dataset was merged with information on referee work during the same period in either UEFA leagues. We also have personal information on referees (age), as well as their past work in our sample (the number of games refereed in previous years).

For each referee-season pair, we define average favoritism bias from referees in two ways. First, for referee r in season s, $Bias_{r,s}$ is the average difference between the stoppage time when the home team is losing vs winning. Second, $Bias\_pred_{r,s}$ is the average deviation from the

[16] See a short discussion in 5.3 in the Appendix.

predicted stoppage time for referee r when the home team is losing vs winning, based on our model (3).

To analyze the relationship between European career and bias, we estimate first a cross-sectional linear probability regression, with the dependent variable, $Euro_{r,s}$ taking up 1 when referee r worked at least a single game at either competitions in season s and 0 otherwise.

Experience is a key potential confounder in this model, and hence we add age (linearly) and dummies for each number of seasons of experience, as well as the number of domestic league games refereed. All right-hand side variables refer to the $s-1$ season. We also add league and season dummies.

$$\begin{aligned} Pr(Euro_{r,s} = 1) = \alpha + \beta_1 Bias_{r,(s-1)} + \gamma_1 Experience_{r,(s-1)} \\ + \gamma_2 Dom\_games_{r,(s-1)} + \gamma_3 Age_{r,s} + L + S + \epsilon_{r,s}, \end{aligned} \tag{6}$$

Alternatives include restricting the sample for the 7th national league season for each referee only (Columns 3,4), replacing bias with predicted bias (Column 2), and having the number of European games instead of a binary variable (Column 4). Finally, (in Column 5) we repeat the model of Column 1, with a different bias definition. $bias\_influence_{r,s}$ is calculated only for influential (top 6 ranked) teams. There are far fewer observations here, so we only estimate a pooled OLS.

Table 3: Referees' favoritism bias and UEFA jobs

	(1)	(2)	(3)	(4)	(5)
Dependent variable	Binary	Binary	Binary	Count	Binary
Estimation	Pooled OLS	Pooled OLS	OLS	OLS	Pooled OLS
Home teams:	All	All	All	All	Influential
Mean home bias	-0.0002		0.0004	-0.0004	-0.0004
	(0.0003)		(0.0010)	(0.0035)	(0.0006)
Mean pred. home bias		-0.0001			
		(0.0004)			
Domestic games (N)	0.0302**	0.0303**	0.0419**	0.2447***	0.135
	(0.0081)	(0.0081)	(0.0172)	(0.0562)	(0.105)
Controls	Yes	Yes	Yes	Yes	Yes
League FE	Yes	Yes	Yes	No	Yes
Season FE	Yes	Yes	No	No	Yes
R^2	0.39	0.39	0.35	0.46	0.39
Observations	574	574	79	79	233

Note: Column 1,2,3,5: Dependent variable is binary, refereed in Europe, linear probability model. Column 4: Count of games in Europe, OLS. In columns 1-4, the bias is the difference in the average residual when the home team is losing vs winning. Predicted bias in column (2) is based on 3. In column 5, it is the difference between influential and non-influential teams when losing at home. Robust standard errors (col 3,4), referee level clustered standard errors (col 1,2,5). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Results are presented in Table 3. We find absolutely no correlation between bias and probability (or count) of refereeing European games. Thus, we see no correlation between any type of favoritism bias from referees and the likelihood of the referee working in the UEFA competitions. Referees do not (or cannot) expect any benefit in terms of European success to help teams that are likely to play in European competitions.

4 Summary

In summary, we found that referees (i) support the home team by allowing the game to last longer if it's losing, (ii) do it even in closed games, (iii) but mostly for influential (highly ranked) teams, and (iv) seemingly unrelated to career benefits. Thus, instead of adherence to social pressure (crowds), we find evidence for adherence to host organizations, especially influential ones. So why do referees support these teams? After ruling out direct social pressure or corruption, we are left with a more nuanced explanation: it may be the consequence of an unconscious bias driven by a persuasive home team organization. People are often uncertain about their decisions and this is when this bias kicks in: favoring or being more willing to err on the side of a persuasive group. For instance, referees may make errors in not compensating enough for wasted time. Such an error may get unpunished if it hurts a small team but may get massive coverage in the media when a top team suffers.

5 Appendix

5.1 Football Rules

5.1.1 Stoppage Time

The length of the stoppage time is carefully but not very strictly defined in professional football. According to Point 7.3 of *Laws of the Game*, the official rules of football maintained by the International Football Association Board, the stoppage time at the end of both halves should compensate for the time lost through substitutions, injuries, wasting time, disciplinary sanctions, cooling breaks, VAR checks, and any other cause such as goal celebrations. Thus, the referee decision is driven by a set of detailed rules regarding which events may generate an extension, but the exact length is determined by the referee[17].

Figure 2 plots the kernel density estimate of stoppage time distribution across all matches in our sample. We can observe spikes around each minute, the largest at 3 minutes.

5.1.2 VAR

The Video Assistant Referee (VAR) system, aims to minimize human errors and their influence on match outcomes. Video replays of key events of the game (such as goals, penalties and red card fouls) are reviewed by an official who communicates with the referee on the pitch. If a potential referee mistake is identified, the game is interrupted for an on-field review of the situation, often lasting several minutes. Time spent reviewing decisions is intended to be compensated for by adding more stoppage time.

5.1.3 Leagues

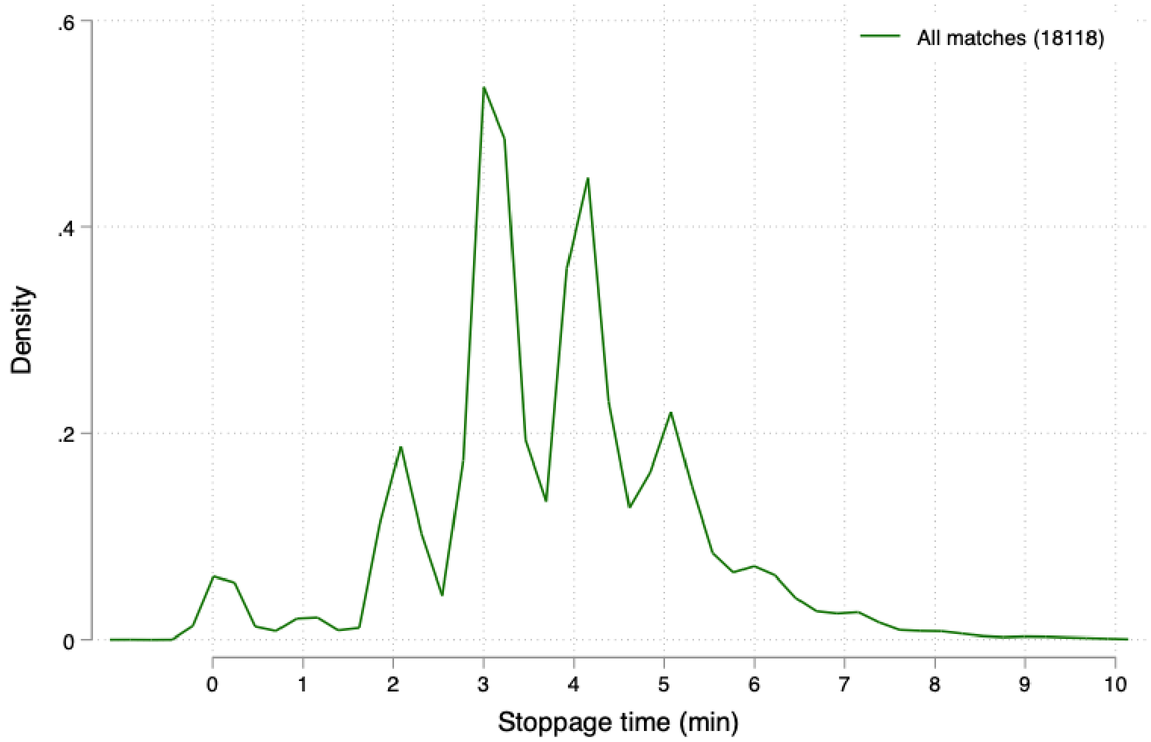
In a sport season, teams play in their national leagues twice with every other team, once home and once as visitor. A win yields three points, a draw one, and the loser gets nothing. These points sum up at the end of season to yield a final league table and create the ranking of teams. The first four teams will play in the UEFA Champions League, the fifth and sixth will play in the less lucrative UEFA Europa League next season. The last two three teams will be relegated to second divisions. A single point will often decide about winning, European spots or relegation.

5.2 Covid and Closures

Due to the Covid-19 outbreak in Europe in Spring 2020, practically every football league was suspended as of the second weekend of March. The last round before suspension was played behind closed doors in Italy on the 8th and 9th of March, as well as the last game in our sample, played on the 11th of March in Germany. The remaining games of the 19/20 season were played, with a more intensive schedule, starting as of the 16th of May in Germany, the 11th of June

[17]See, for example, the following quote by former referee Dermot Gallagher: "*(...)we've had this standardisation that we're going to play 30 seconds per substitution, and for excessive goal celebrations we're to play another 30 seconds – so it starts to tot up, and this is why we find the three or four minutes we have on average at most games.*" – https://playtheadvantage.com/2014/05/27/how-stoppage-time-is-determined/

Figure 2: Distribution of additional time



in Spain, the 17th of June in England, and as of the 21st of June in England. Each of these games was played behind closed doors, with no fans allowed to enter the stadium. In France, the remaining games of the 19/20 season were not played.

Season 20/21 started in August 2020 in France and in September 2020 for the rest of the leagues. The vast majority of the games were played behind closed doors. Depending on the severity of the Covid situation, some leagues allowed a restricted number of fans to be present in the stadium for short periods throughout the season. This was the case for France between August and October 2020; for England and Germany between September and October 2020, and in May 2021; for Italy and Spain in May 2021. As Table 4 shows, this partial opening meant that the stadiums were filled to 10-15% of capacity on average. No match was played with full capacity of fans during season 20/21 – the highest attendance-to capacity ratio in our sample is 34%.

Table 4: Attendance of matches during Covid

Season	N	N closed	N open	Mean attendance	Max attendance
England 19/20	380	92	0		
England 20/21	380	346	34	12.3	25.9
France 19/20	279	1	0		
France 20/21	378	316	62	15.1	33.8
Germany 19/20	306	83	0		
Germany 20/21	306	269	37	10.4	20.5
Italy 19/20	380	132	0		
Italy 20/21	380	379	1	1.3	1.3
Spain 19/20	380	111	0		
Spain 20/21	378	373	5	11.9	21.7

Note: The last two columns indicate the mean and maximum, respectively, of attendance-to-capacity percentage among games where reported attendance was greater than zero.

5.3 Big Team Influence at UEFA

Formally, teams can have no say over referee jobs.[18] However, there were several news reports when teams tried to exert pressure on UEFA regarding referees.[19]

Another area where teams do have influence is the format of European games. Indeed, as early as 1997, major teams achieved to make the Champions League include runner up from 8 leagues, only to be extended to the current 4 teams from top 5 leagues. In 2021, UEFA proposed changes in the Champion League to please the big teams that were trying to force the Super league: https://www.mirror.co.uk/sport/football/champions-league-plan-premier-league-26923957, indeed the new Champion League format (after the Super League proposal) benefits the big teams (mainly English ones): https://www.forbes.com/sites/steveprice/2022/05/11/uefas-champions-league-changes-benefit-the-big-six-and-newcastle-united/?sh=44446ce57743

There are several cases exposed by Football Leaks that, according to UEFA regulations, would end up with certain penalties that were not applied in the end, such as for Manchester City and PSG https://www.dailymail.co.uk/sport/sportsnews/article-6347551/Leaked-documents-claim-Manchester-City-hid-30m-UEFA-FFP-investigators.html

There are other tax and doping cases where UEFA helped cover up https://www.spiegel.de/international/world/football-leaks-doping-tests-and-real-madrid-a-1240035.html

5.4 Data Cleaning

For a few games, the source of our event data contains either no information or obviously erroneous information, such as extremely few events recorded. As we cannot construct the

[18] See https://documents.uefa.com/r/Regulations-of-the-UEFA-Champions-League-2022/23/Article-48-Appointment-and-replacement-of-referees-Online.

[19] For example, see https://www.firstpost.com/sports/champions-league-juventus-blames-uefa-chief-refereeing-officer-pierluigi-collina-of-bias-against-serie-a-clubs-4429027.html.

measures of interest for these matches, our analysis excludes them.

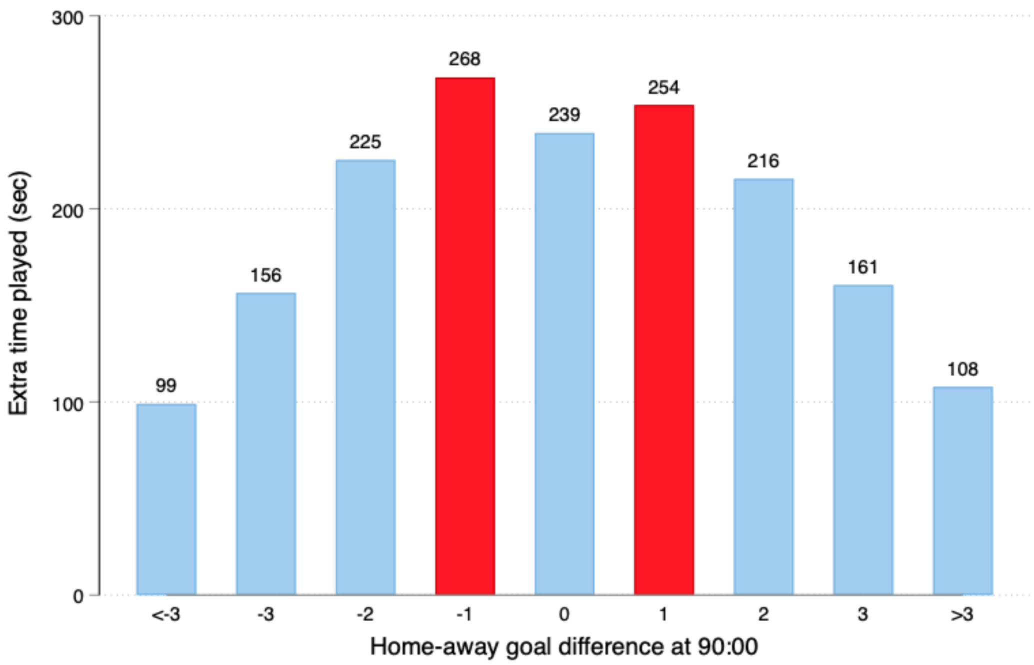
We excluded one game with more than 33 minutes of wasted time corresponding to a severe head injury of the goalkeeper, which involves on-field medical care and, thus, a long period of play stop.

Our *Losing offensiveness* measure is not observed for 7 games in our sample, implying that on these games the losing team did not have a single pass in the stoppage time. Thus, we impute a value of 0 for these in our analysis.

5.5 Descriptive Statistics

In the first graph, we reproduce Figure 1 of (Garicano et al., 2005), showing the average stoppage time by goal difference at 90:00. As shown by Figure 3, we see a rather similar pattern: the average stoppage time is longer for tighter matches. Since goals in stoppage time are rare events, matches with more than one goal difference at the end of the regular time are highly likely to be already settled, and in this case referees tend to blow the final whistle significantly earlier, serving the interest of both teams.

Figure 3: Stoppage time awarded by score margin



The next table presents some informative descriptive statistics about key events, team values and more.

Figure 4 looks at densities only for games, where the home team is winning or losing by a single goal, and we can see that for shorter extensions, we have more games of home team winning, while for longer one the opposite is true. A Kolmogorov-Smirnov test also confirms the statistical significance of the difference between the two distributions.

Table 5: Summary statistics

	Mean	SD	Median	Min	Max	N
Stoppage time (sec)	260.03	77.42	248	3	1065	7021
Home losing (0/1)	0.44	0.50	0	0	1	7021
Wasted time (sec)	1125.66	177.95	1118	523	2004	7021
Cards	3.13	1.87	3	0	13	7021
Subs	5.75	1.15	6	2	10	7021
Fouls	14.38	4.17	14	3	33	7021
Goals	1.40	1.13	1	0	7	7021
Losing offensiveness	60.96	9.03	61.52	7.10	100.00	7014
Goals in stoppage time	0.16	0.39	0	0	3	7021
Round	19.15	10.78	19	1	38	7021
Home odds	2.76	1.95	2.20	1.05	23.00	7017
Away odds	4.37	3.38	3.40	1.10	41.00	7017
Odds difference (H-A)	-1.60	4.63	-1.20	-39.93	21.90	7017
Home rank	10.65	5.56	11	1	20	7021
Away rank	10.25	5.63	10	1	20	7021
Rank difference (H-A)	0.40	7.95	1	-19	19	7021
Home team value (m EUR)	8.43	9.58	4.81	0.81	64.14	7021
Away team value (m EUR)	9.10	10.22	5.07	0.81	64.14	7021
Value difference (H-A)	-0.67	12.73	-0.17	-61.60	61.60	7021
VAR (0/1)	0.32	0.47	0	0	1	7021

Note: The sample consists of games with a one goal difference at 90:00, from seasons between 2011/12 and 2020/21 of the top 5 European football leagues. [GH2] denotes variables describing the second half of the games only. See Section 2.2 for a detailed description of the variables.

5.6 Alternative Model Estimations

In this subsection, we present additional model estimations and robustness checks.

First, Table 6 provides more details regarding variables for Table 7. It indicates how each of the control variables is related to the length of stoppage time. We see that all factors have a strongly significant effect on stoppage time.

Next, Table 7 shows that our baseline estimations are not sensitive to the choice of fixed effects. This includes fixed effects for the referees.

5.7 Comparison to Other Relevant Papers

In this subsection, we show that not methodological difference may explain deviation from earlier results.

Table 8 replicates the main models of Garicano et al. (2005) and we can see that the estimated coefficient remains much smaller than theirs. In particular, Table 9 argues that the attendance to capacity ratio is to a large extent explained by the home and away teams themselves, along by the league and the season. As such, no proper inference can be made on the basis of the variation of this measure.

Figure 4: Comparing the cumulative distribution of home team leading vs losing

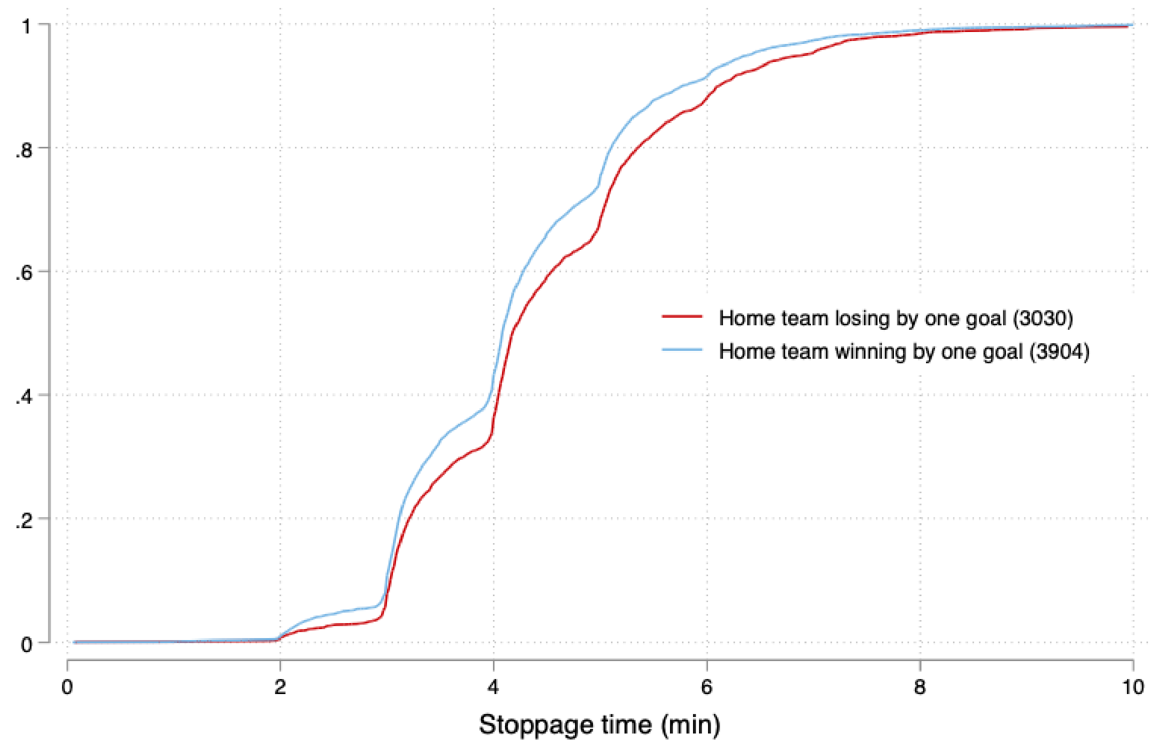


Table 6: Regressions indicating home-team bias

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
Home lose	14.28***	13.38***	11.10***	12.80***	12.72***
	(2.24)	(2.04)	(1.79)	(1.79)	(1.76)
Wasted time		0.20***	0.18***	0.19***	0.19***
		(0.01)	(0.01)	(0.01)	(0.01)
Cards		3.32***	4.70***	4.46***	4.32***
		(0.61)	(0.47)	(0.47)	(0.47)
Subs		2.94***	7.10***	6.72***	7.22***
		(0.83)	(0.86)	(0.83)	(0.79)
Fouls		-4.05***	-2.50***	-2.52***	-2.29***
		(0.37)	(0.24)	(0.25)	(0.25)
Goals		-4.92***	-4.53***	-4.91***	-4.62***
		(0.81)	(0.75)	(0.78)	(0.82)
Losing offensiveness		0.24***	0.32***	0.35***	0.32***
		(0.09)	(0.08)	(0.08)	(0.07)
ET goals		21.89***	21.01***	20.83***	20.90***
		(2.62)	(2.38)	(2.37)	(2.30)
VAR		21.39***	24.07***	24.49***	27.11***
		(2.82)	(2.23)	(2.24)	(2.55)
Round		-0.15**	-0.22***	-0.23***	-0.23***
		(0.07)	(0.06)	(0.07)	(0.07)
Constant	253.77***	46.66***			
	(2.61)	(9.63)			
League FE	No	No	Yes	Yes	Yes
Home team FE	No	No	No	Yes	Yes
Referee FE	No	No	No	No	Yes
Observations	7021	7021	7021	7021	7001
R^2	0.01	0.28	0.38	0.41	0.44

Note: Standard errors clustered at home team level.

Table 7: Regressions indicating home-team bias

	(1)	(2)	(3)	(4)	(5)
Home lose	14.28***	13.38***	11.10***	12.80***	12.72***
	(2.24)	(2.04)	(1.79)	(1.79)	(1.76)
Controls	No	Yes	Yes	Yes	Yes
League FE	No	No	Yes	Yes	Yes
Home team FE	No	No	No	Yes	Yes
Referee FE	No	No	No	No	Yes
Observations	7021	7021	7021	7021	7001
R^2	0.01	0.28	0.38	0.41	0.44

Note: Standard errors clustered at home team level. Controls include time with ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, round of season, whether VAR was used, and losing offensiveness (average distance of the passes of the losing team from opponent's goal line in the stoppage time). Note that 20 referees worked on a single game only and hence, the last column has 20 observations less.

Table 8: Garicano replication

	(1)	(2)	(3)	(4)	(5)
Home lose	11.12***	10.72***	8.11**	6.08	12.83*
	(1.94)	(1.97)	(3.49)	(3.80)	(6.86)
Yellow	7.66***	7.65***	7.74***	8.22***	8.21***
	(0.52)	(0.54)	(0.53)	(0.56)	(0.55)
Red	20.86***	21.70***	21.33***		
	(2.04)	(2.08)	(2.02)		
Subs	14.90***	14.87***	15.10***		
	(1.27)	(1.34)	(1.29)		
Home value million	-0.82**	-0.78**	-0.38***	-0.72***	-0.69***
	(0.32)	(0.32)	(0.14)	(0.16)	(0.16)
Away value million	-0.46***	-0.44***	-0.48***	-0.62***	-0.65***
	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)
Home rank	0.15	0.13	-0.38	-0.39	-0.38
	(0.32)	(0.32)	(0.26)	(0.28)	(0.28)
Home-away rank diff	0.21	0.17	0.18	0.33*	0.31
	(0.18)	(0.17)	(0.17)	(0.19)	(0.19)
Round			-0.27***	-0.22***	-0.23***
			(0.09)	(0.07)	(0.07)
Home lose × Round			0.15		
			(0.15)		
Attendance 1000				270.09***	190.12*
				(89.50)	(96.43)
Home lose × Attendance 1000				172.04	253.34*
				(125.37)	(134.82)
Attendance/Capacity (%)					0.16**
					(0.08)
Home lose × Attendance/Capacity (%)					-0.12
					(0.09)
League FE	Yes	Yes	Yes	Yes	Yes
Referee FE	No	Yes	No	No	No
Season FE	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	No	No	No
R^2	0.30	0.34	0.27	0.24	0.24
Model	T2C4	T2C6	T5C4	T6C3	T6C4
Observations	6151	6129	6151	5654	5654

Note: Standard errors clustered at home team level. Controls include time with the ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, whether VAR was used, and the average distance of the passes of the losing team from opponent's goal line in the stoppage time.

Table 9: Regressions explaining attendance/capacity ratio

	(1)	(2)	(3)
Away favorite	10.15***	5.05***	7.64***
	(0.77)	(0.71)	(0.76)
Season FE	Yes	Yes	Yes
League FE	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes
R^2	0.76	0.76	0.76
Favorite definition	bigteams	oddsdiff	valuediff
Observations	16840	6524	6524

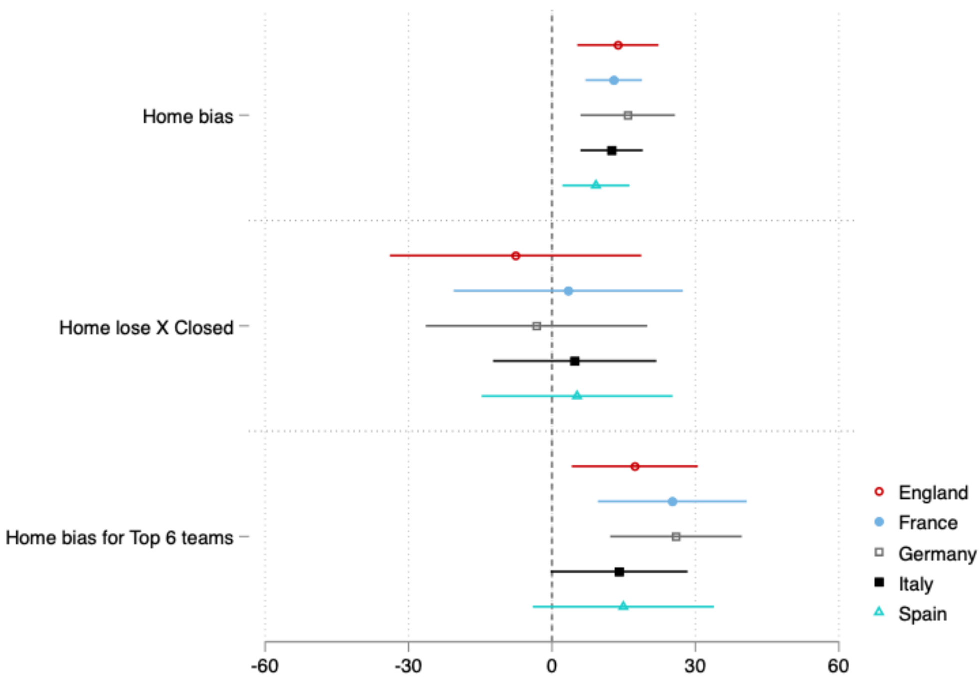
Note: Standard errors clustered at home team level. Model 1 includes all matches in our sample, while Model 2 and 3 concerns only those included in our main analysis: matches with a 1-goal difference at the end of regular time.

5.8 Persistent Results across Leagues and Time

Different countries and leagues may have different customs and regulations. Thus, we might see heterogeneity for any of our results, or find that they are driven by peculiarities in a single country.

To illustrate how heterogeneous our main findings are across countries, we run regressions of Models (2) and (4) of Table 1 for each league separately (without league fixed effects). The estimated coefficients are presented in the Appendix Tables 10, and 11.

Figure 5: Estimated regression coefficients by league



We find that all main results are highly robust across leagues. First, the home-team bias is very similar, ranging between 9.2 and 15.8 seconds (not statistically different from each other). Second, we see that during closed games, all leagues experienced a small change only, with point estimates ranging between -7 and +5 seconds, neither being statistically different from zero. Third, in terms of the moderator variable of influential teams, the interaction term of the home lose indicator and the indicator for the top 6 teams is rather stable across leagues, ranging between 10-17 seconds (neither is statistically different). Note, however, that broken down by leagues, this result lacks statistical power, and the results are not significant. This is because top teams rarely lose at home, and the number of observations by league is too small. Figure 5 summarizes our main findings.

Tables 10 and 11 show the regressions visualized on Figure 5.

Table 10: Regressions by league

	(1)	(2)	(3)	(4)	(5)
Home lose	13.79***	12.91***	15.84***	12.47***	9.19**
	(4.33)	(3.01)	(5.03)	(3.32)	(3.57)
Home lose × Closed	-7.61	3.40	-3.25	4.74	5.25
	(13.42)	(12.23)	(11.82)	(8.72)	(10.19)
Controls	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes	Yes
Observations	1467	1486	1086	1503	1479
R^2	0.35	0.39	0.41	0.34	0.38
League	England	France	Germany	Italy	Spain

Note: Standard errors clustered at home team level. Controls include time with the ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, whether VAR was used, and the average distance of the passes of the losing team from opponent's goal line in the stoppage time.

Table 11: Regressions by league

	(1)	(2)	(3)	(4)	(5)
Home lose	10.70**	11.13***	13.37***	10.50***	6.27
	(4.51)	(3.71)	(4.69)	(3.33)	(4.15)
Home lose × Home Top 6	9.56	11.20	10.75	10.94*	16.77**
	(10.57)	(7.23)	(8.41)	(5.91)	(7.17)
Controls	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes	Yes
R^2	0.35	0.39	0.42	0.35	0.39
League	England	France	Germany	Italy	Spain
Observations	1467	1486	1086	1503	1479

Note: Standard errors clustered at home team level. Controls include time with the ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, whether VAR was used, and the average distance of the passes of the losing team from opponent's goal line in the stoppage time. All these variables are also interacted with the *Closed* dummy.

5.9 Alternative Influential (Top-ranked) Team Definitions and Results

In our main specification to measure influence, we used team ranking by achievement. There are several alternative measures to this ranking.

First, perceived quality may be better better captured with **pre-match odds.** To measure perceived quality, we use the pre-match odds, cleaned from incorporated home advantage. We achieve this by estimating $\widehat{Odds}_{ij} = \alpha + \beta Home_{ij} + \epsilon_{ij}$ where $Home_{ij}$ is an indicator whether team i is the home team in match j. Then, for each match j with a one-goal difference at 90:00, we take the home-away difference between the estimated ϵ_{ij}s. This metric is used only for difference between teams. Pre-match odds come from of the betting site bet365.com for home win, draw, and away win.[20]

Second, monetary wealth may be better proxied by the **estimated squad value** of each team at the start of the season. This is based on adding up individual player values for squads using estimates from Transfermarkt. For example, in the 2018/19 season, Arsenal, an English Premier League team is valued at 659 million euros, the 6th most valued team and making the 5th position in the points table. https://www.transfermarkt.com/premier-league/startseite/wettbewerb/GB1/plus/?saison_id=2018.

Third, another monetary measure is revenues. The teams with **Top revenues over 10 years** is a very simple metric could be considering the top 20 teams that have generated the most revenue in the last 10 years as favorite of the matches they are playing, unless these top teams play against each other. The latter matches, along with those involving no top team, are labelled as matches with no clear favorite. It is a binary variable by design. The revenues data was collected from Deloitte Football Money League Wikipedia page. The top 20 teams are the following: Arsenal, Chelsea, Liverpool, Manchester City, Manchester United, Tottenham from England; Lyon, Marseille, Paris Saint-Germain from France; Bayern Munich, Borussia Dortmund, Hamburg, Schalke 04 from Germany; AC Milan, Inter Milan, Juventus, Roma from Italy; Atletico Madrid, Barcelona, Real Madrid from Spain.

In Table 12 below we reproduced key results with alternative influential team definitions.

Table 12: Robustness home-team bias heterogeneity

	(1)	(2)	(3)
Home lose	11.87***	9.31***	10.60***
	(1.80)	(1.74)	(1.79)
Home favorite	1.01	-4.13	-2.66
	(4.48)	(2.73)	(2.53)
Home lose × Home favorite	11.09**	18.97***	12.28***
	(5.11)	(4.18)	(4.40)
Controls	Yes	Yes	Yes
League FE	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes
R^2	0.41	0.42	0.41
Favorite definition	20 richest	Odds diff	Squad value diff
Observations	7021	7021	7021

Note: Games with a single goal difference after regular time, N=7021. Dependent variable is stoppage time in seconds. Controls include time with ball out of play, number of cards, substitutions, fouls, goals, goals in stoppage time, round of season, whether VAR was used, and the average distance of the passes of the losing team from opponent's goal line in the stoppage time. Control variables are also interacted with the *closed* dummy. Standard errors clustered at home team level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

[20]The betting data was downloaded from football-data.co.uk, and was available for all but nine matches of our event data sample.

References

Austen-Smith, D., and Fryer, R. G. (2005, May). An economic analysis of 'acting white'. *Quarterly Journal of Economics*, *120*(2), 551–583.

Behn, D., Berge, T. L., and Langford, M. (2018). Poor states or poor governance? explaining outcomes in investment treaty arbitration. *Northwestern Journal of International Law and Business*, *38*(3), 1.

Bursztyn, L., and Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, *9*, 131–153. doi: 10.1146/annurev-economics-063016-103625

Carmichael, F., and Thomas, D. (2005). Home-Field Effect and Team Performance: Evidence From English Premiership Football. *Journal of Sports Economics*, *6*(3), 264–281. doi: 10.1177/1527002504266154

Dawson, P., and Dobson, S. (2010). The influence of social pressure and nationality on individual decisions: Evidence from the behaviour of referees. *Journal of Economic Psychology*, *31*(2), 181–191.

DellaVigna, S., and Gentzkow, M. (2010, September). Persuasion: Empirical Evidence. *Annual Review of Economics*, *2*(1), 643-669.

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003, 08). Persuasion Bias, Social Influence, and Unidimensional Opinions*. *The Quarterly Journal of Economics*, *118*(3), 909-968.

Dohmen, T. (2008). The influence of social forces: Evidence from the behavior of football referees. *Economic inquiry*, *46*(3), 411–424.

Dohmen, T., and Sauermann, J. (2016). Referee Bias. *Journal of Economic Surveys*, *30*(4), 679–695. doi: 10.1111/joes.12106

Duggan, M., and Levitt, S. D. (2002, December). Winning isn't everything: Corruption in sumo wrestling. *American Economic Review*, *92*(5), 1594-1605.

Erikstad, M. K., and Johansen, B. T. (2020). Referee bias in professional football: Favoritism toward successful teams in potential penalty situations. *Frontiers in Sports and Active Living*, *2*.

Garicano, L., Palacios-Huerta, I., and Prendergast, C. (2005). Favoritism under social pressure. *Review of Economics and Statistics*, *87*(2), 208–216. doi: 10.1162/0034653053970267

Gauriot, R., and Page, L. (2019, October). Fooled by performance randomness: Overrewarding luck. *The Review of Economics and Statistics*, *101*(4), 658–666.

Hodler, R., and Raschky, P. A. (2014, 03). Regional Favoritism *. *The Quarterly Journal of Economics*, *129*(2), 995-1033.

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

Parsons, C. A., Sulaeman, J., Yates, M. C., and Hamermesh, D. S. (2011, June). Strike three: Discrimination, incentives, and evaluation. *American Economic Review*, *101*(4), 1410–1435.

Prendergast, C., and Topel, R. H. (1996). Favoritism in organizations. *Journal of Political Economy*, *104*(5), 958-978.

Schwartzstein, J., and Sunderam, A. (2021, January). Using models to persuade. *American Economic Review*, *111*(1), 276-323.

Vollaard, B., and van Ours, J. C. (2022). Bias in expert product reviews. *Journal of Economic Behavior and Organization*, *202*, 105-118.