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, v1.11 – 10 August 2022. See latest version HERE.

Abstract

This paper investigates potential mechanisms of social pressure leading to biased decisions. We revisit and extend Garicano, Palacios-Huerta, and Prendergast (2005) who found that football referees bias the length of stoppage time in favor of the home team. They argue that home bias arises in an attempt to satisfy the crowd in the stadium. Using exogenous variation in crowd size, along with a substantially finer and wider dataset, we propose a new mechanism – persuasion by host organizations. By including events with empty stadia, we show that the bias is uncorrelated with home crowd. Instead, we find that the bias mostly comes from favoring big and successful teams of the leagues. We find no evidence, however, of biased behaviour leading to better career outcomes for the football referees in Europe. This suggests that referees are not so much affected by social pressure from crowds or corruption but persuasion by the host team organization.

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

JEL-codes: D71, Z20, C21

1 Introduction

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. It has broad consequences in a wide range of situations ranging from court cases to allocation of new stocks at IPOs. One source of favoritism is social pressure, the exertion of influence by a person or group. According to the *American Psychological Association*, channels of social pressure are persuasion (informational influence), calls for conformity (normative influence), and direct influence (VandenBos, 2007). Conformity is about agents' desire to conform to some expectations in hope of future rewards or in adherence to a social image (Bursztyn and Jensen, 2017). Persuasion,

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

where individuals or groups may try influence agent's decision to adhere to their interest via persuasive communication (DellaVigna and Gentzkow, 2010). Direct influence or even corruption happens when an agent is offered or may reasonably expect future gains (or avoid losses) if adhering to a certain behaviour.

Despite its importance, empirical evidence about various channels is limited. As argued by DellaVigna (2009), it is difficult to distinguish social and psychological motives when it comes to observational data. Moreover, in line with Bursztyn and Jensen (2017), individuals may even care about how they are perceived by more than one reference groups, or may be influenced by more groups at once. As a possible solution, social scientists turned to sports data to gain insight about behavior of agents in a variety of settings driven by available data and some exogenous variation in rules. For instance, expectation of financial rewards were shown to lead to match rigging in Japanese Sumo wrestlers (Duggan and Levitt, 2002). Another prime example of such work is Garicano et al. (2005) who analyzed decisions by association football (soccer) referees who were shown to favor the home team. The paper has been frequently quoted being a rare case of identifying favoritism in the form of biased decision owing to influence of outside groups (fans in stadium).

Our paper revisits the football referee setting with a focus to deconstruct the mechanism behind the referees' biased decisions, and to offer evidence about the persuasion channel. To that end, we complied a novel and comprehensive dataset on European football games covering match events, referee decisions, and referee career paths, and relied on an an exogenous variation in the size of the outside groups owing to stadium closures due to the Covid pandemic in 2020 and 2021. The new data and the setup allowed us to disentangle social pressure by fans and possible persuasion activities by host organizations to argue for the role of the latter in explaining biased decisions. Furthermore, we rule out the existence of direct pressure (or corruption) channel with an analysis of referee career paths shown to be unrelated to favoritism biases.

Our starting point is Garicano et al. (2005) who found that referees use discretionary power to systematically award more stoppage time to the home team when it stands to lose at the end of regular time, thus allowing more chances to catch up. The discretionary power allows referees to set the amount of stoppage time allowed at the end of the second half of the game to compensate for the time lost for a variety of events such as injuries. They argue that the mechanism behind this bias is that referees internalize in their decisions the preferences of the home team's supporting crowd in the stadium. 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 yield lower bias. These findings are extended by Dohmen (2008) who suggested that distance to crowds – in stadiums with running tracks – exacerbated the influence of crowds.

A key difficulty in such analyses is that the crowd size (and the type of stadium) is endogenous to the quality and wealth of host organization. There is a very strong correlation among financial background, history, sport success and amenities: richer teams tend to be more successful playing in larger and better stadiums. This correlation makes it hard to tell channels of social pressure apart.

First, we need data on more leagues to filter out possible league specific rules and customs and ensure that any result may have high external validity. Large coverage is also necessary to power our identification. We also need very detailed data that allows partialling out game characteristics and records information about referees as well. To meet these demands, we created a comprehensive, novel, fine granularity dataset on 10 seasons (2011/12 to 2020/21) of the 5 most prestigious leagues in Europe: the English *Premier League*, the Spanish *La Liga*, the Italian *Serie A*, the German *Bundesliga 1*, and the French *Ligue 1*. 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.

Comparing narrow margin standings after regular time, we find that referees on average add 13 seconds more stoppage time when the home team is losing compared to the case when the home team is winning. This is somewhat smaller than the 20 second bias that Dohmen (2008) found 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 two Spanish seasons of 1994/95 and 1998/99.

Second, to tell apart crowd size and other features of home advantage, our key contribution is to use stadium closures as a natural experiment to disentangle the mechanism behind the home-team bias in referees' stoppage time decisions. In particular, we use the Covid-19 pandemic as an exogenous variation in stadium attendance to assess whether the bias can be attributed to fans present in the stadium. We show that the bias is unchanged even without spectators.

This result suggests that other form of influence must be in play: persuasion stemming from proximity to the host team and its organization. Here hospitality may play an important role: on match day, referees are surrounded by members of the host organization often spending the whole day at premises. Before, during, and after the match, the home team staff accounts for the safety and well-being of the referee, so he has some incentives to please them or at least not to make them hostile. Gifts from the home team are not uncommon either.[1]

If influence by team organizations matter, we shall expect greater bias when the influence is stronger. Indeed, we show that exposure to more successful (in terms if performance) teams is associated with a larger bias. Referees add more than twice as many seconds for the top 6 teams compared to the rest[2]. Note that our coverage of multiple leagues and seasons is necessary to test the power of the crowd effect during Covid.

Third, we investigated if persuasion may itself be confounded by direct pressure or corruption. While direct bribery is extremely unlikely[3], referees may expect professional rewards in the future when helping big teams [4]. To refute this alternative hypothesis, we examined referees' careers in the prestigious European competitions where only the top teams play. If favoritism were driven by expectations of more work in these leagues, referees with higher bias would be picked more –

[1]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`.

[2]The idea of top team bias in football has been shown in a different setting. Using Norwegian data and an expert panel, (Erikstad and Johansen, 2020) showed that the top 2 teams are more likely to get a penalty awarded.

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

[4]Big teams are able to pressure UEFA to change competition format such as the Champions League that has grown to include more and more big clubs from top leagues or avoid penalties, for details, see 5.8.

in collusion with participating teams. We find no such evidence (neither in domestic leagues).

How can we reconcile our results with earlier ones? We suspect that earlier results on crowds may have been confounded by team characteristics that remain in play even in closed doors (or become even more salient): more successful teams have larger stadiums, and less likely to have running tracks (no big team in Germany has tracks, while 9.8% of others have tracks).

Taken together, we empirically measure a new channel of social pressure to prompt favoritism: persuasion by powerful groups, in a setting where it may be told apart from crowds or corruption. Instead of adherence to social pressure by crowds, we find evidence for adherence to host organizations, especially powerful ones. Interestingly, referees may yield to pressure by powerful organizations even if they don't expect career gains. Note that our setting is very special: it is an extremely open and transparent situation. Thus, in more opaque settings, we expect a larger bias. This finding may have even greater external validity: think about independent advisors, accountants, consultants working for organizations – their advice may be biased towards their hosts, especially for large firms even in the lack of bribery.

The paper is broadly related to an emerging literature on using sports data to analyze human behavior.[5] More narrowly, the paper is related to earlier research, mostly in sport economics and psychology about referee bias for home teams for other outcomes such disciplinary action (yellow and red cards, free kicks, or penalties).[6] However, disciplinary decisions are always intertwined with the actions of the players making it difficult to empirically disentangle referee behavior from player behavior in these cases. Indeed several studies analyzed how the game may have changed during Covid.[7] The relative length of stoppage time allotted by the referee, however, does provide us with a measurable, isolated variable of referee behavior.

In what follows, we first describe the dataset we used in Section 2. Then we discuss our results in Section 3 in several steps: present evidence on home-team bias in referee decisions, show that it is not driven by the crowd in the stadium, speculate about possible channels and argue that an important factor is favoritism to successful teams. Finally, we also show the stability of our results over time and across countries. Section 4 concludes.

2 Data and empirical strategy

In this section, we first present our dataset, composed of several data tables from several sources. Second, we describe our core model and the variables we used. Third, we explain how we relied on closed stadiums due the covid pandemic as natural experiment.

[5] Gauriot and Page (2019) also use football data to talk about quality perception biased by luck, while Parsons, Sulaeman, Yates, and Hamermesh (2011) looks at discrimination in US baseball.

[6] For referee's home team bias, see e.g. Carmichael and Thomas (2005); Dawson and Dobson (2010), and for a review, Dohmen and Sauermann (2016); for more recent evidence on games during Covid see Bryson, Dolton, Reade, Schreyer, and Singleton (2021); Scoppa (2021); Reade, Schreyer, and Singleton (*forthcoming*).

[7] One example is Caselli, Falco, and Mattera (2022) who found evidence of African players doing better during closed doors.

2.1 Data sources

We use a sample of matches from 10 seasons (from 2011/12 to 2020/21) of the top 5 European leagues (England, France, Germany, Italy, Spain). In the German league, there are 18 teams each season; in the rest of the leagues there are 20. Each team hosts every other team once in a season in its own stadium. This coverage would imply 18,260 games.[8] Our actual sample is slightly smaller. First, due to the Covid-19 outbreak, season 2019/20 of the French Ligue 1 finished prematurely, with only 279 out of 380 games played. Due to some deficiencies in our main data sources, we lost another 41 games. Thus, our final sample consists of 18,118 games.

Our dataset was created using several data sources.[9] First, we use an an 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, and corner. In a typical game, there are around 750 events per team recorded before stoppage time, and around 30 events per team recorded during stoppage time, corresponding to around 17 events per minute, or an event being tracked every third or fourth second on average.

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 leagues. The data has been collected from `soccerway.com`.

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

2.2 Empirical strategy

Our main outcome variable of interest is *Stoppage_time* is the length of playtime (measured in seconds) beyond the regular time (90:00 minutes). It is calculated from the the event data based on the timestamp of the event indicating the end of the game. It may be slightly different than the stoppage time added, as referees may add seconds based on events like fouls and substitutions in the stoppage time.

We follow (Garicano et al., 2005), and focus only on the cleanest comparison to study referee bias: looking at matches where the goal difference at the end of 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 = 7021 games (out of 18,118).

For empirical exercises, events are aggregated at the game and half level; regular time and stoppage time separately. Since referees compensate for the wasted time at the end of both halves, we only consider events taking place in the second half of the game for each of these

[8] $10 \times 4 \times 20 \times 19 \ + 10 \times 1 \times 18 \times 17 = 18260$
[9] Data and code will be made available.

variables. The average stoppage time is around 4.5 minutes, ranging from 3 seconds to almost 18 minutes on the two extremes. 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 4 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 by the home teams. To partial these out, we estimate the following model with OLS:

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

where our unit of observation is a single game played between home team h, and away team a in season s in league l. 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 ($Controls_{h,a,s}$).

First, we have created a proxy measure for the amount of wasted time by counting the time during which it is reasonable to assume that the ball was out of play. We did this by adding up the seconds passing 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. Using this variable (*Wasted time*, measured in seconds) we are able to control for the effective playing time in the second half of the match and, as such, the best available approximation of the justifiable length of stoppage time. Our estimated wasted time in the second half varies between 9 minutes and more than 33 minutes. The average wasted time is 19 minutes (out of 45+4), which is equivalent to about 60% effective playing time, in line with the general findings of the sports press[10].

Second, we also control for the number of yellow or red cards, substitutions, fouls, and goals in the second half. These events are associated with the longest play stops. Although in theory our wasted time measure should capture all the effects of these variables, we include them in our specifications for two main reasons. First, instead of precisely calculating the seconds with the ball out of play, referees tend to use heuristics such as the number of these key events when they make their decisions on the stoppage time.[11] Second, 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. Having said that, all our results are robust to leaving out some or all of these factors from the models.

Third, a potential confounding effect may be that instead of favoring the home team, referees simply let the attacks started during the end of the stoppage time finish, even beyond the indicated length. It is a confounder because in general, the losing team is likely more offensive during stoppage time (because they need to score 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

[10] See e.g. https://football-observatory.com/IMG/pdf/mr64en.pdf.

[11] 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/

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.

Fourth, during the examined 10-year period, a technological innovation has been introduced that may also affect the mechanisms underlying the bias of our interest. Both the use of VAR system and the compensation for the time lost may depend on whether the home or the away team is winning. To control for this potential confounder, we generate a league-season level variable (VAR) indicating whether the technology was in use[12].

Fifth, we also control for the round of the season for each match. 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). The higher this number, the more stake there is in the game. It is reasonable to assume that the pattern of the matches changes as the season advances and by controlling for this factor, we rule out a potential confounder.

Finally, we also include league and home team fixed effects δ_l and η_h. 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. The choice of fixed effects set, however, does not affect our results.[13]

3 Results

This section presents our empirical findings. We confirm the existence of a referee bias towards the home team, and show that this bias is not driven by the supporters in the stadium. As a next step, we show that influential teams enjoy larger referee bias. 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 crowds

In our dataset, the raw difference between the additional time extended if the home team is losing compared to when the home team is winning by a single goal margin is 13 seconds. This remains unchanged once all control variables are added. For more details about the distribution of stoppage time, and how controls affect the estimated coefficient, see Section 5.2 in the Appendix[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

[12]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 5.2

[13]See Table 6 in Appendix.

[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 are a magnitude smaller than the one in Garicano et al. (2005). In lack of available data from the period it is hard to make a direct comparison. We estimated their core model in their Table 2, column 4, and found a similar estimate to our favored specification. It is possible that referee behavior in the nineties in Spain had less oversight indeed.

Table 1: Regressions indicating home-team bias

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs
Home lose	13.10***	12.80***	13.22***	13.17***
	(1.77)	(1.79)	(1.77)	(1.77)
Home lose × Covid			-1.35	
			(5.46)	
Home lose × Closed				-0.71
				(5.44)
Controls	Yes	Yes	Yes	Yes
League FE	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes
R^2	0.42	0.41	0.41	0.41
Period	pre-Covid	all	all	all
Observations	6151	7021	7021	7021

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 the average distance of the passes of the losing team from opponent goal in the extra time. All these variables are also interacted with the Covid or Closed dummy.

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%. As the Table 5 in the Appendix shows, in these two years, about 2/3 of games were behind fully or partially closed stadiums. 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:

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

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 referees may take different amount of time to make decisions depending on broad circumstances such as closed stadium.

The natural experiment, induced by the outbreak of the Covid-19 pandemic, allows us to test the hypothesis that the referee bias favoring the home team is the consequence of the social pressure exerted by the supporters of the home team in the stadium. Table **??** shows that eliminating crowd from the stadium has no effect on referee bias. The estimated interaction terms are zero in all specifications: with and without home team FE, and using all Covid games or fully closed ones.

This result suggests that while referees are as biased as before in supporting the home team when it is losing, but their actions cannot be driven by crowd size. Even in the extreme and exogenous case of closed games, the home-team bias remains as it was before. Thus, crowds cannot be the main driver.

This results deviates from established wisdom on the role of crowds generating social pressure. It is, however, very robust to specifications and also across leagues – see Figure 4 in Section 5.6 in the Appendix.

3.2 The home-team bias driven by influential teams

In this section we investigate if team influence is indeed correlated with the size of the home-team bias. Influence is the capacity to exert social pressure. Team influence is hard to measure directly. It may stem from amenities, wealth of the club, quality of the team (squad, management, the whole organization) as shown by performance in the season. In the case of football teams, influence and success are very close concepts.

In this context, we think about team influence as a function of two concepts: wealth and quality. Wealth could be measured as the value of current players as well as wealth in a broader sense taking into account past achievements and brand value. Quality could be measured by the total points or the ranking in the league, pre-game odds or Elo rating, a longer run measure of teams' strength. Our baseline specification will use league table ranking as it captures the influence of teams in an easy way. Ranking is defined as the end-of-season ranking 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 in a year is close to expected average ranking throughout the year.[15] Nevertheless, all of these wealth and quality measures are highly correlated with a correlation coefficient between 57% and 92%.

Note that influence is likely to be non-linear as it also based on the ability to play in the two pan-European competitions: typically the top 6 teams will play in either the UEFA Champions League or in the UEFA Europa League. Thus, we will also consider a binary variable to measure influence being inside or outside the top 6 teams per league.

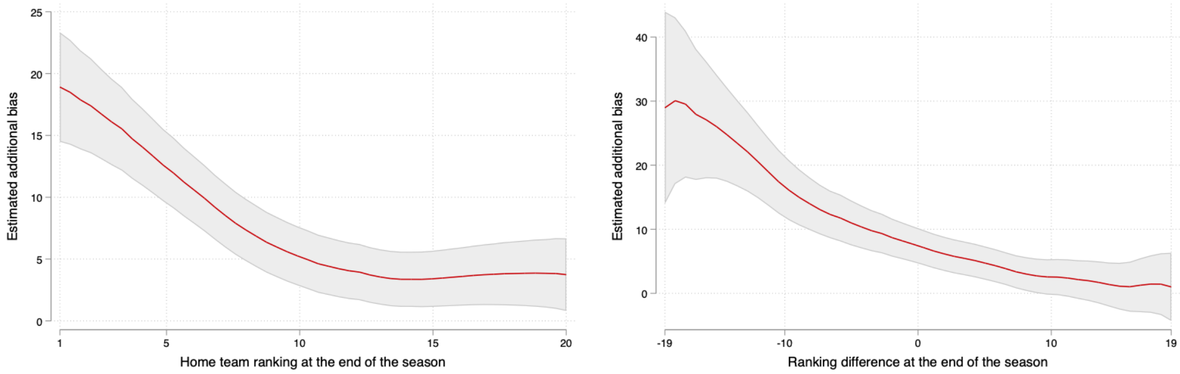
To better understand the relationship between stoppage time and ranking, we first estimate a model with only the football rule controls (such as time with ball out of play, number of cards, substitutions), as described in Section 3.1. Then, we compute the difference between predicted and actual stoppage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

$$Bias = Stoppage\_time_{h,a,s} - \hat{\gamma}Controls_{h,a,s} + \hat{\delta}_{league} + \hat{\eta}_{hometeam_{h,s}} \tag{3}$$

In the second step, using local polynomial smoothing regressions, we plot this bias against

[15] The advantage of this measure is that it is what teams care about, an easy metric to interpret, can be compared easily across leagues. Unlike past performance and all teams have a value for it in all seasons.

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



(a) Home team ranking

(b) Difference between home and away team ranking

Note: 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 goal in the stoppage time) all interacted with *Closed* dummy.

heterogeneity by the ranking of the home team. In Figure 1a, we see a fairly strong pattern with better teams enjoying a greater home-team bias. This is especially true for the top few teams. In Figure 1b, we see a similar pattern with the difference plotted against bias: when the top teams play against lower ranked ones, the gap is 30 seconds, but it goes down to 0 when minnows play at home versus against the top teams.[16]

We estimate this heterogeneity in two ways. First, we consider variation by the home team influence, with β_3 measuring the heterogeneity in home-team bias. The $Home\_ranking_{h,s}$ variable may be estimated in a linear form or with a binary indicator variable (influential teams defined as the top 6 teams in each league).

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

Second, an alternative model uses the difference between home and away team, 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 the difference is defined as a gap greater than -10 (defined in terms of season-end league position).

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

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

The results of both models are presented in Table 2. The simplest setup is considering the binary Top 6 ranked team indicator in column (1) of Table 2. Influential (top-ranked) teams enjoy a home-team bias that is more than twice the one for non-influential ones (10.23 vs 22.98 seconds).

Looking at the home team's ranking in a linear way in Column (2), we see 19.8 seconds for the number 1 team ($20.48 - 1 \times 0.67$), and this gain is reduced by 0.67 seconds per rank. For the last team in the league, the home-team bias shrinks to 7.1 seconds (20.48-20 $\times$ 0.67).

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, we once again find that when there is a sizeable gap in favor of the home team, the bias is twice the size compared to non such difference (12.05 vs 27.70 seconds). In the linear model, for equal teams the bias is 14 seconds and the slope is -0.84. In case of a difference of 19 ranks, the estimated bias is -2 seconds, $13.94 - 0.84 \times 19$, and in the other direction, it is 29.9 seconds.

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.

Note that as the different aspects of success (such as points and ranking, financial power, size of supporting base, clout outside the league, etc.) in football are strongly correlated, we cannot empirically dissect whether the effect comes, in general, through money or popularity. Indeed, we find that our results are qualitatively robust to the two alternative measures (such as odds difference and squad value difference) as well.[17]

Finally we can look at how the difference between the influential (top ranked) team and others may have changed during closed games. In column (5), we repeated column (1) now interacted with the closed period. We found that in these games, influential teams are the only teams that are helped actually (37.38 seconds versus almost none for non influential ones). This points to a situation where the host organizations play a very important role (protection against Covid), hence seem to have an even greater influence on referees. This is another source of evidence on pressure coming from the organizations.

To summarize, we find that influential (top-ranked) teams benefit substantially more from referee decisions. The result is stable over time, and if anything, the gap in benefit has further increased in closed games. When playing at home the top influential teams seem to create an

[17]See the coefficient table for these in Appendix (Tables **??** and **??**.

Table 2: Regressions indicating home-team bias heterogeneity

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
Home lose	10.23***	20.48***	12.05***	13.94***	11.18***
	(1.84)	(4.41)	(1.76)	(1.86)	(1.79)
Home lose × Home Top 6	12.75***				9.11***
	(3.74)				(3.34)
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)	
Home lose × Closed					-8.22
					(6.39)
Home lose × Home Top 6 × Closed					28.28**
					(10.93)
Controls	Yes	Yes	Yes	Yes	Yes
League FE	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes	Yes
R^2	0.41	0.41	0.41	0.42	0.42
Observations	7021	7021	7021	7021	7021

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 the average distance of the passes of the losing team from opponent goal in the stoppage time. All these variables are also interacted with the *closed* dummy. Column (5) reports the estimated value of the key variables interacted with the *closed* dummy.

environment that makes referees more inclined to help. It is the influence of the host team organization and not the crowd size that determines referee behavior.

This finding is important because it may explain earlier evidence. Team influence is correlated with many observable characteristics: stadium attendance and capacity, even the type of stadium teams have. For instance, home team, league, and season, along with an indicator of a top away team explain more than 75% of the variation in the attendance rate across all matches in all samples; see Table 12 in the Appendix. Similarly for running tracks, the top 6 teams in Germany have no tracks in their stadium but of the rest do. Thus, the hypothesized effect can be more reliably identified on the extensive rather than on the intensive margin, in particular by focusing on how the bias changed when no fans were present in the stadiums at all.

3.3 Is there corruption? Referee career in Europe

Finally, we investigate if career concerns may motivate referees to help influential teams. The main question we address here is whether more biased referees are more likely to get work at the UEFA Champions League and in the UEFA Europa League games. Working in these European games are a pinnacle of both player and referee careers. Influential teams with positions at UEFA boards can block unwanted referees and possibly promote preferred ones[18].

To analyze careers, we aggregated game information into an unbalanced panel of referees ($r = 1 \dots 233$) and seasons ($s = 2011 - 12 \dots 2020 - 21$). This panel ($N = 1148$) covers 1-10 seasons for 233 referees. On average we observe a referee for 4.9 seasons in their national leagues work on 70 games, 33 of which would have a one goal difference after 90 minutes. For each referee and season, we define average referee bias in two ways. First, $Bias_{r,s}$ is the average seasonal difference between the stoppage time for referee r when the home team is losing minus the time when the home team is winning. Second, $Bias\_pred_{r,s}$ is the average seasonal (s) deviation from the predicted stoppage time for referee r when the home team is losing minus the average deviation when the home team is winning, based on our model (3).

This new dataset is referee * season level, but only contains referees who had both at least one 0:1 and 1:0 result after regular time result in a season. Dropping observations based on a single game, we are left with 179 referees and $N = 711$ referee*season observations. This dataset was merged with information on referee work during the same period in the Champions League and in the Europa League. We also have information some key personal features of the referees (age), as well as the past of referees in our sample (number of games refereed in previous years).

To analyze the relationship between European career and bias, we estimate first a cross-sectional linear probability regression, with the dependent variable, $euro_{r,s}$ is one when referee r worked at least a single game at either competitions in season s and zero otherwise. Experience is a key potential confounder in this model, and hence we add league experience dummies. We also add referee age, and number of national league games. All right hand side variables refer to the $s - 1$ season.

$$Pr(euro_{r,s} = 1) = alpha + \beta_1 bias_{r,(s-1)} + \beta_2 experience_{r,(s-1)} + \gamma_1 age_{r,s} + \epsilon_{r,s}, \tag{6}$$

Alternatives include restricting the sample for the 7th national league season for each referee only (column 3,4), replacing bias with predicted bias (column 2), and having the number of European games instead of a binary variable (column 4).

Results are presented in Table 3. We find absolutely no correlation between bias and probability (or count) of European games.

Finally, we repeat both regressions with a different bias definition, $bias_influence_{r,s}$ - the same difference but for only for influential (top 6 ranked) teams – yielding the same negative result. There are much fewer observations here, so we only estimate a pooled OLS with season dummies as well as a variable counting the number of season a referee worked in the domestic league. As shown in Table 3 column (5), once again, we find no correlation.

[18]See for example XXXXX

Table 3: Referee 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	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
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: Linear probability model regressions (column 1,2,3,5), Robust standard errors (col 3,4), referee level clustered standard errors (col 1,2,5). 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.

Thus, we see no correlation between any type of referee bias and the likelihood of the referee working in the UEFA competitions – a key metric of career success in European football.

4 Discussion

In this paper, we extended and revisited earlier evidence of favoritism under social pressure, measured by home-team favoring bias among European football referees. Using extensive data consisting of more than 7 thousand football matches from 10 seasons of the 5 top European leagues, we find that referees support the home team by allowing the game to last longer if they are losing. Importantly, the bias is not due to crowds as it exists even in the absence of home fans, but it is larger when favors an influential (top-ranked) teams, and is especially large when an influential team is losing at home to a minnow. We showed that referee bias is, however, uncorrelated with career benefits.

Why do referees support successful teams? After ruling out direct social pressure or corruption, we are left with some more nuanced explanations. 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. Maybe under an influence or s they are more willing to err on the side of persuasive group. For instance, risk aversion happens as referees make errors including that they do not compensate 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. The more successful (wealthy, star-studded) the host, the greater this influence.

[One closing para on why this is important]

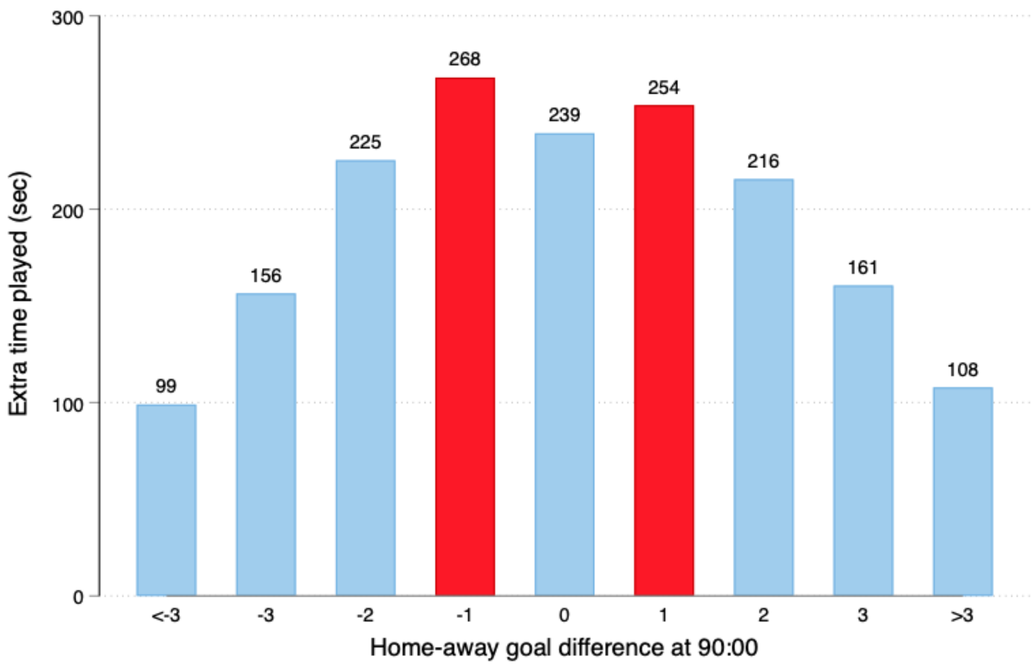
5 Appendix

5.1 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. Similarly to that table, in our data, shown by Figure 2, the average stoppage time is the longer the tighter the match is. 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.[19]

The gap between home or away team losing is rather different to ours. Also, more observations might have led to a more balanced distribution as well.

Figure 2: stoppage time awarded by score margin



The next table presents some informative descriptive statistics.

5.2 Football rules: stoppage time, VAR

The length of the stoppage time is 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. Towards the end of the regular play

[19]Because a draw (that is, zero goal difference) may be useful for one or the other team depending on many aspects of the game and the season, we do not base our analysis on drawing games.

Table 4: 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. All variables concern events happening in the second half of the games only (where this is relevant). See Section 2.2 for a detailed description of the variables.

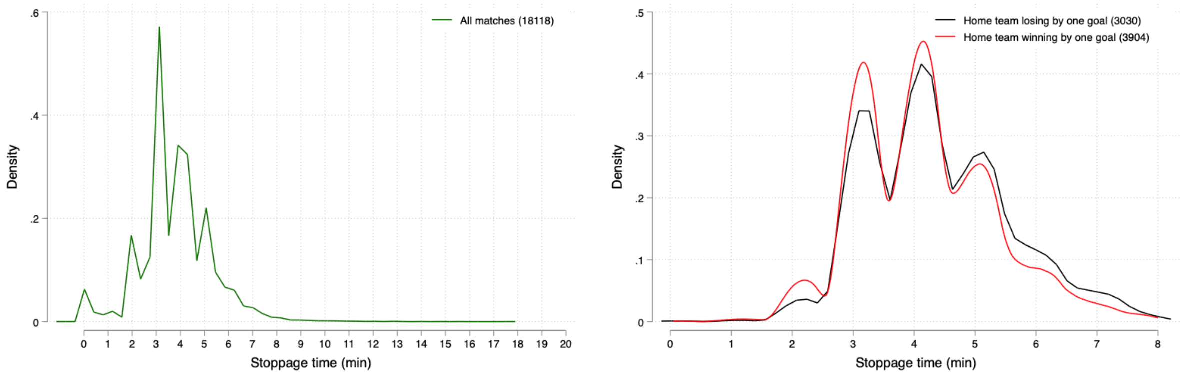
time, the fourth official indicates the minimum additional time. The actual stoppage time can be increased but not reduced by the referee.

The referee decision is driven by a set of detailed rules regarding which events may generate an extension. The exact length is, however, determined by the referee.

Figure 3 plots the kernel density estimate of stoppage time distribution for these two outcomes separately (Panel 3b), along with that for the full sample (Panel 3a). We can observe spikes around each minute: the largest at 3 minutes in the full sample. We can also see that as we move towards longer stoppage time, home losing matches are dominating starting the fifth additional minute. A Kolmogorov-Smirnov test also confirms the statistical significance of the difference between the two distributions.

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.

Figure 3: Distribution of additional time



(a) All matches

(b) Matches with one goal difference

5.3 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 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 supporters 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 supporters 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 5 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%.

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

We excluded a 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

Table 5: 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.

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 Alternative model estimations

Table 6 shows that our baseline estimations are not sensitive to the choice of fixed effects.

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)
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 goal in the stoppage time). Note that 20 referees worked on a single game only and hence, the last column has 20 observations less.

Table 7 presents the models of Table 6 with indicating how each of our control variables is related to the length of stoppage time. We see that all factors have a strongly significant effect on the outcome variable.

Table 7: 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 8 replicates the main models of Garicano et al. (2005). See the specific table and column reference in the Table.

Tables 9, 11 and 10 show the regressions visualized on Figure 4.

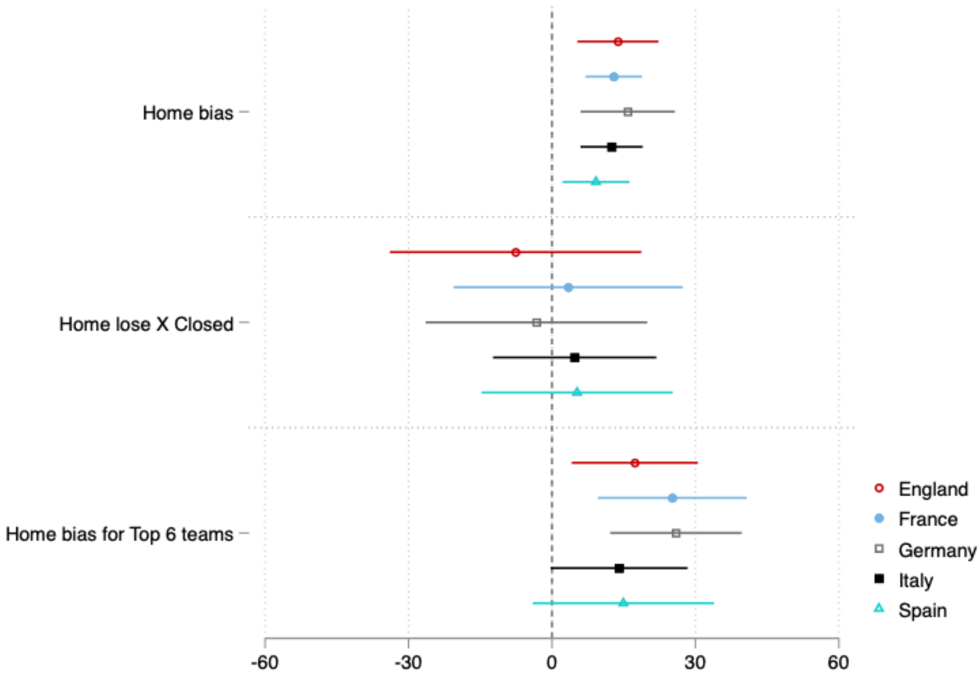
Table 12 argues that the attendance to capacity ratio is to a large extent explained by the home and away teams themselves, along by the league and the season. As such, no proper inference can be made on the basis of the variation of this measure.

5.6 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 Model (4) of Table **??** and Model (1) of Table 2 for each league separately (without league fixed effects). The estimated coefficients are presented in the Appendix Tables 9, and 10.

Figure 4: 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 4 summarizes our main findings.

5.7 Alternative influential (top-ranked) team definitions and results

There are several alternative measures to ranking used.

Estimated squad value of each team at the start of the season. This is based on adding up individual player values for squads.

Top 20 in 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.

Regarding financial power, data was collected from this 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.

Points in the league is the end-of-season number of points per team, the higher the better. Points are typically between 20 and 100. The ranking in the actual year is best predictor of perceived quality, but an alternative is the past year value.

Pre-match odds. To measure 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]

Elo ranking is a ranking based on the cumulative quality of opponents. In our estimation, values are demeaned at the league level.

Below, we reproduced key results with alternative influential team definitions.

[To be added.]

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

Table 8: Garicano replication

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
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 goal in the stoppage time.

Table 9: Regressions by league

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
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 goal in the stoppage time.

Table 10: Regressions by league

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
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 goal in the stoppage time. All these variables are also interacted with the *Closed* dummy.

Table 11: Regressions by league

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
Home lose	12.87***	12.97***	16.45***	15.32***	12.18**
	(4.13)	(3.78)	(4.62)	(2.62)	(4.54)
Home lose × Home-away rank difference	-0.61	0.22	-1.56***	-1.82***	-0.91**
	(0.45)	(0.46)	(0.50)	(0.38)	(0.34)
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 goal in the stoppage time. All these variables are also interacted with the *Closed* dummy.

Table 12: Regressions explaining attendance/capacity ratio

	(1) Attendance/Capacity (%)	(2) Attendance/Capacity (%)	(3) Attendance/Capacity (%)
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 Big team influence at UEFA

Formally, teams can have no say over referee jobs[21]. However, there were several news reports when teams tried to exert pressure on UEFA regarding referees[22].

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

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

[22] 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

References

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

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

Caselli, M., Falco, P., and Mattera, G. (2022, Forthcoming). When the stadium goes silent: How crowds affect the performance of discriminated groups. *Journal of Labor Economics*.

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. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, *47*(2), 315–72.

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

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.

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.

Reade, J., Schreyer, D., and Singleton, C. (*forthcoming*). Eliminating supportive crowds reduces referee bias. *Economic Inquiry*.

Scoppa, V. (2021). Social pressure in the stadiums: Do agents change behavior without crowd support? *Journal of Economic Psychology*, *82*(August 2020), 102344. Retrieved from https://doi.org/10.1016/j.joep.2020.102344 doi: 10.1016/j.joep.2020.102344

VandenBos, G. R. (2007). *Apa dictionary of psychology.* American Psychological Association.