Received: 23 December 2022 | Accepted: 12 June 2024

DOI: 10.1111/ecin.13245

ORIGINAL ARTICLE

Economic Inquiry

Favoritism under multiple sources of social pressure

Gábor Békés[1,2,3] | Endre Borza[2,4] | Márton Fleck[1]

[1]Central European University, Vienna, Austria

[2]HUN-REN Centre for Economic and Regional Studies, Budapest, Hungary

[3]C.E.P.R., London, UK

[4]Center for Collective Learning, CIAS, Corvinus University, Budapest, Hungary

Correspondence
Gábor Békés, Central European University, Quellenstrasse 51, Vienna 1100, Austria.
Email: bekesg@gmail.com

Abstract

When social pressure leads to favoritism, policies might aim to reduce the bias by affecting its source. This paper shows that multiple sources may be present and telling them apart is important. We build a novel and granular dataset on European football games and revisit the view that supporting crowds make referees help the host team. We find this bias to remain unchanged even in stadiums closed due to Covid-19. Instead, influential host organizations emerge as the source of social pressure. This has an adverse effect on maintaining the ranking of influential teams and hindering the progress of smaller teams.

KEYWORDS

corruption, favoritism, persuasion, social pressure, sport

JEL CLASSIFICATION

D71, Z20, C21

1 | INTRODUCTION

In competitive situations between organizations, an agent (such as a judge or a referee) often has the opportunity to make a decision benefiting one party. Favoritism occurs when a decision-maker gives preferential treatment to one party at the expense of another, without justification from rightful determinants. Favoritism appears in various contexts, from promotion decisions in organizations (Prendergast & Topel, 1996) to the allocation of regional development funds (Hodler & Raschky, 2014).

Beyond explicit corruption (bribery), another common source of favoritism is social pressure: the exertion of influence by a person or a group. Social pressure includes conformity and persuasion. Conformity involves agents' desire to conform to expectations, seeking approval or adhering to a social image of themselves (Bursztyn & Jensen, 2017). Persuasion occurs when individuals or groups influence an agent's decision in their interest via persuasive behavior such as communication (DellaVigna & Gentzkow, 2010), looking for both cognitive and emotional responses from the receiver (DeMarzo et al., 2003; Schwartzstein & Sunderam, 2021).

Abbreviations: NBA, National Basketball Association; UEFA, Union of European Football Associations; VAR, video assistant referee.

Managing Editor: Rob Simmons

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 straightforward when there is a single channel of social pressure. However, individuals may care about the perception of more than one reference group (Bursztyn & Jensen, 2017). For instance, students may care about both the perceptions of their peers and those of their prospective employers (Austen-Smith & Fryer, 2005). Regarding persuasion, consumers might be targeted by various parties simultaneously.

We contribute to the literature by showing the importance of distinguishing sources of social pressure. We revisit Garicano et al. (2005), who analyzed the decisions of association football (soccer) referees and found evidence of favoritism, specifically referees favoring the home team in crucial situations.[1] They argued that this bias is driven by social pressure from the home team supporters in the stadium. We argue that referees may be influenced by multiple sources of social pressure, such as the supporting crowd and the host team organization. Using exogenous variation in crowd size due to stadium closures during the Covid-19 pandemic, we find no effect from crowds on referee behavior. Instead, we show that the host team's organization influences referees. In particular, more successful teams have greater influence and benefit more from referees' favoritism bias. Nevertheless, social pressure remains the driver of favoritism, as we find no evidence of corruption: the bias in favor of successful teams is uncorrelated with referee career paths.

To differentiate these sources, we need appropriate data and an identification strategy. We compile a comprehensive dataset on European football games covering match events, referee decisions, and referee career paths. Our high granularity dataset spans 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 actions, penalties, and injuries) along with a timestamp.

The data covers a long period before the Covid-19 pandemic as well as games played during the pandemic.

This paper confirms the existence of favoritism as shown by Garicano et al. (2005): referees use their discretionary power to systematically award more stoppage time to the home team when it stands to benefit.[2] Comparing games with a one-goal difference after regular time, we find that referees, on average, add 13 s more stoppage time when the home team is losing compared to when the home team is winning.[3]

Next, we explore the sources of social pressure and examine the role of supporting crowds. Garicano et al. (2005) argue that referees' home bias is driven by social pressure from home team supporters in large numbers in the stadium: referees conform and internalize the preferences of home team supporters. This mechanism is identified via variations in crowd size and its composition: larger attendance generates greater bias, while a higher attendance-to-capacity ratio—indicative of a higher share of visiting crowds—yields lower bias. These findings are extended by Dohmen (2008), who suggested that the distance from 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-19 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 the relative size of home supporters, we use stadium closures as a natural experiment. Using this exogenous variation in the 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 investigate. The organization hosts the 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. On match day, 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 looks after the safety and well-being of the referee. Gifts from the home team are not uncommon either.[4]

At the same time, referees in close proximity to players and managers could be the strongest channel. Players, especially stars, will confront referees on decisions they do not like. Famous players and managers will exert star power. Both the persuasion by these individuals and the human desire to conform to greatness could play a role.[5] Furthermore, players, when ex-ante expected to win at home, could exert higher levels of pressure on the referees when losing or drawing as they themselves are feeling the pressure.[6]

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 referees, the desire to conform to successful teams and their star players will be stronger. These teams will also mobilize more resources (better facilities and personnel) for persuasion.[7] To detect social pressure related to the host team organization, we compare teams of different rankings: more successful teams are expected to enjoy a larger home team bias.

Looking at the difference in stoppage time when it benefits the home team compared to when it benefits the away team, we indeed find that referees add more than twice as much stoppage time at the end of the second half for the benefit of the most successful third of teams compared to the rest.[8] 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 to show an example where social pressure comes from different sources. We found that the traditionally assumed source, namely 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 the 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 the rest have tracks). A large and highly granular dataset and exogenous variation were necessary to distinguish these two sources of social pressure.

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

Finally, we examine whether a more obvious explanation for favoritism could also act as a key possible confounder: corruption. While direct bribery is extremely unlikely,[9] referees may expect professional rewards in the future when helping the most successful teams.[10] 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 often—in collusion with participating teams. We find no such evidence.

Several previous studies have looked at sports outcomes in games played behind closed doors, without supporting crowds. Bryson et al. (2021), Reade et al. (2022), Wolaver and Magee (2022), and Morita and Araki (2022) all found that in terms of disciplinary action, the home bias changed substantially during closed games. However, while such actions are easier to observe, identification is problematic, as these referee decisions are intertwined with the actions of the players. This makes it difficult to empirically disentangle referee and player behavior.[11] Importantly, Caselli et al. (2022) identified a differentiated effect of crowd presence on player behavior and performance. Specifically, African players performed better in closed games. This indicates that support or opposition of attending crowds can have diverse and specific impacts on individuals participating in the games.

Another related paper is Gong (2022), who assesses whether referees' home bias changed without crowd support in NBA games. He compares the probabilities of incorrect referee decisions made against home versus away players in the last 2 minutes of narrow-margin games. He finds that the empty arenas due to the Covid-19 pandemic have no differential effect on home bias. However, in this basketball example, no baseline home bias was identified even before closed games. This means that the confounders of the referees' bias, such as team influence, could not be investigated either.

This paper contributes to understanding that social pressure may be related to several actors simultaneously, with different policy outcomes. Consider possibly biased arbitration judgments in investor-state dispute settlements (Behn et al., 2018). There may be various social pressures: persuasion by and conformity to wealthy countries and crowd pressure via social media or protests. In the case of investigating bias in online reviews (Vollaard & van Ours, 2022), reviewers may be subject to social pressure when favoring big brands popular among large sections of consumers, as well as be targets of persuasion by companies sending gifts and offering marketing events. In both these scenarios, our results suggest that analysis must tackle different sources rather than assuming any. The paper is also related to the literature on multi-sender communication (Battaglini, 2002; Gentzkow & Kamenica, 2016), showing a case of strongly correlated signals by the senders.

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

In what follows, we first describe the dataset we used and the key empirical methods we applied 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. 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 in the top five European leagues (England, France, Germany, Italy, and 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.[12]

Such a dataset has several 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. Likewise, 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. This records each and every action happening on the pitch. Each event has a type and a timestamp (at the second level). Event types include: pass, ball recovery, foul, tackle, throw-in, free-kick, yellow and red card, substitution, penalties, shot, goal, and corner. In a typical game, an event happens once every 3.6 s: there are 1432 events recorded in the regular playtime of 90 min, and 90.1 events during stoppage time. Where relevant, the dataset also contains the location of the event on the pitch (in terms of x and y coordinates).

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 have been collected from soccerway.com.

Furthermore, we used complementary data from a variety of sources. Information on stadiums comes from transfermarkt.com as well as from Wikipedia pages of teams.

We used Deloitte Football Money League to identify top clubs.[13] We collected data to estimate the squad value of each team for every season, using historical player valuations from transfermarkt.com. Finally, for each league and season, we downloaded the clubs' Elo rating score at the start of the season from elorating.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 min). We measure this as the timestamp of the game end in the event data. This is set to indicate the time on the clock (beyond 90 min) when the referee blows the final whistle. Note also that due to technical reasons, a typical gap of 2–3 s arises between the clock shown for the actual whistle and the end of the match timestamp in our data. However, as shown in Appendix A1.2, this is orthogonal to the regular time standing and will hence not affect our results.[14]

The measure of *Stoppage_time* can (and often does) slightly differ from the expected stoppage time indicated by the referee at the end of the regular time, as referees may adjust the indicated time based on events like fouls and substitutions that happen during the stoppage time. Again, this variable is measured in seconds, allowing us to measure changes at a high level of precision.

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 regular time (90:00) is exactly one goal. Our key independent 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). 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 253 s, ranging from 3 to 660 s. The home team wins slightly more games than the away team (57% compared to 43%), consistent with the well-documented general home advantage in sport (Jamieson, 2010). For a broader review of descriptive statistics, see Table A2 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 the playing style of the home team. To partial these out, we estimate the following model with OLS:

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

where our unit of observation is a single game played between home team h, and away team a in season s 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 ($Controls_{h,a,s}$) as described below. Standard errors are clustered at the home team level.[15]

To adjust for factors affecting stoppage time, we include the following variables. First and most importantly, we approximate the justifiable length of stoppage time by counting the time during which the ball is likely to have been out of play in the second half. Our granular event dataset introduced above allows us to calculate the measure of *Wasted time* as 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. Thus, this variable captures all the events associated with wasting time, including those that determine the length of the stoppage time as per the Laws of The Game (see Section A1 of Appendix). To make sure that the measure also captures longer interruptions of the games such as injuries or cooling breaks, we also add any interruption of the game that is longer than 30 s. In addition, to make sure that our results are not driven by extraordinary games with very long interruptions (such as a serious injury), we exclude matches that fall within the top 5% of relevant matches in terms of the longest interruption.[16] In the second half, wasted time varies between 11 and 33 min, its average is 21 min (out of 45 + 4.5), equivalent to c. 60% 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 further confounding effect may be that instead of favoring the home team, referees may simply let the attacks that 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.

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.[17]

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 at stake in the game, which can affect both the referee's and team's behavior.

Finally, in line with earlier literature, we add league and home team fixed effects (θ_l and η_h, respectively) to capture the footballing style and quality of the team. This allows within-team comparisons of stoppage time conditional on the end of regular time results. 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. That said, all our results are robust to including referee fixed effects, as well as to including referee age (as a proxy for their experience) as a control variable.

3 | RESULTS

This section presents our empirical findings. We confirm the existence of a referee favoritism bias toward 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.

3.1 | Home-team bias is still there, but not because of the crowd

In our dataset, 12.4 s 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 13.41.[18] Detailed regression results are shown in Table A3 in the Appendix,

TABLE 1 Presence of home bias, no crowd effect.

	(1)	(2)	(3)	(4)
Home lose	12.40***	13.41***	13.81***	13.70***
	(2.07)	(1.65)	(1.60)	(1.61)
Home lose × Covid			−2.41	
			(5.19)	
Home lose × Closed				−0.99
				(5.08)
Controls	No	Yes	Yes	Yes
League FE	No	Yes	Yes	Yes
Home team FE	No	Yes	Yes	Yes
R^2	0.01	0.45	0.46	0.45
Observations	6667	6667	6667	6667

Note: Games with a single goal difference after regular time. 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.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

revealing how each of the control variables is related to the length of stoppage time in our baseline specification with various fixed effects (Model 2 of Table 1). We see that stoppage time is indeed correlated with events in the game such as fouls or disciplinary actions. Furthermore, it shows that our baseline estimations are not sensitive to the choice of fixed effects, including fixed effects for the referees.

Due to the Covid-19 outbreak in Europe in the Spring of 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 the 2020/21 season and in spring 2021, while partial opening meant that for 139 games stadiums were filled to an average capacity of 13%. Over these 2 years, about 2/3 of the games were played in fully or partially closed stadiums (For details, see Appendix Table A1 in the Appendix.). We created two indicator variables, *Closed* = 1 when attendance is zero, and *Covid* = 1 which also includes very low attendance games in 2021. As partial opening only affects 6.6% of the 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 Equation (1) and estimated:

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

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 whether the game is played in full or (mostly) empty stadiums: the estimated interaction terms are very close to zero. This result is robust to changing the mix of control variables (including adding referee fixed effects).[19]

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 the existence of bias in referee decisions, but we can rule out that this is driven by the size of the crowd, as the bias remains unchanged even in the extreme case of closed stadiums. This is our first main result: any difference in the crowd size is very unlikely to be the mechanism behind social pressure, therefore, referees must be helping the home team for other reasons.

3.2 | The home-team bias driven by influential teams

If 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 allow teams to have influence: they attract players, fans, investment, or media interest. Influence will also provide teams with the capacity to exert social pressure in the form of the persuasion of independent agents.

Financial and sporting success are strongly correlated: wealthier teams will have better players and will win more often. 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.

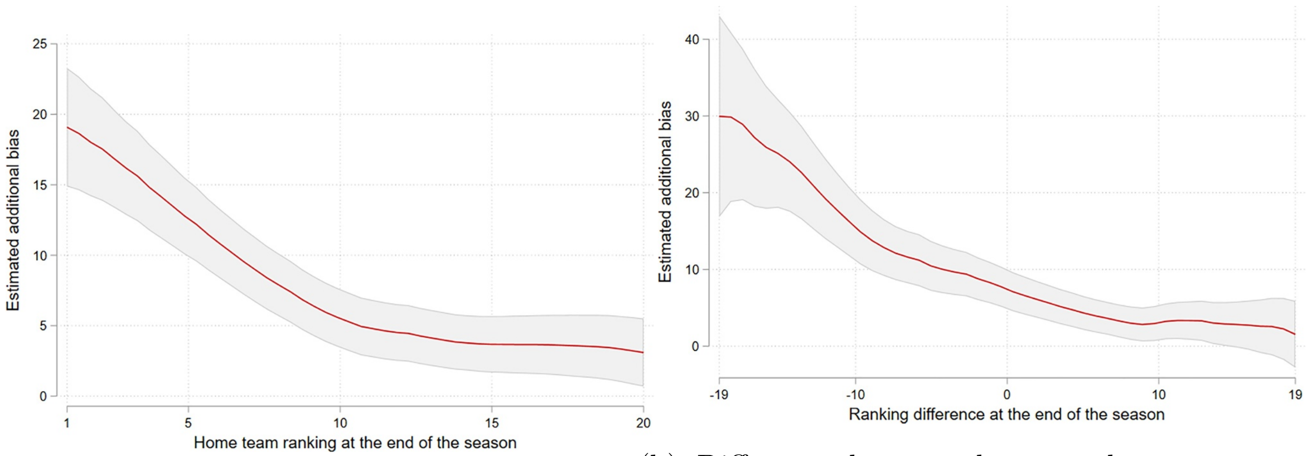
Ranking is defined as the end-of-season position of the team in the league table, the lower the better (1 is the title winner, 18 or 20 is the last team). The final ranking of a team at the end of a season is close to its expected average ranking throughout the season, and it may be easily compared across 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, and substitutions), as described in Section 3.1. Then, we compute the difference between predicted and actual stoppage time. This deviation, $\widehat{Bias}$ from Equation (3), is the measure of the unexplained difference.

$$\widehat{Bias} = Stoppage\_time_{h,a,s} - \left(\hat{\gamma}\, Controls_{h,a,s} + \hat{\theta}_l + \hat{\eta}_h\right) \tag{3}$$

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 s, but it goes down to 0 when a minnow plays at home against a top team.[20]

Both graphs suggest that influence in terms of ranking may be non-linear. This may stem from an important rule: the top six teams will play in the rewarding European championships in the following season. Thus, we will also consider a binary variable, Top6, to measure influence by being inside or outside the top six teams per league.



(a) Home team ranking

(b) Difference between home and away team ranking

FIGURE 1 Heterogeneity of home-team bias by team ranking. (a) Home team ranking. (b) Difference between home and away team ranking. 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, all interacted with the *Closed* dummy.

We estimate this heterogeneity in two ways. First, we investigate the 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.

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

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 a high difference is defined as a gap greater than −10. In both models note that home team fixed effects allow within-team comparisons of different opponents.

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

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 (10.9 vs. 23 s). Looking at the home team's ranking linearly in Column (2), we see 22.08 − 1 × 0.75 = 21.33 s for the top team, reduced by 0.75 s per rank, shrinking to 7 s (22.08 − 20 × 0.75) 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).

TABLE 2 Regressions indicating home-team bias heterogeneity.

	(1)	(2)	(3)	(4)
Home lose	10.86***	22.08***	12.25***	14.29***
	(1.74)	(3.51)	(1.65)	(1.71)
Home lose × Home top 6	12.09***			
	(3.05)			
Home lose × Home rank		−0.75***		
		(0.27)		
Home lose × Home-away rank diff ≤−10			17.35***	
			(5.74)	
Home lose × Home-away rank diff				−0.80***
				(0.20)
Controls	Yes	Yes	Yes	Yes
League FE	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes
R^2	0.46	0.46	0.46	0.46
Observations	6667	6667	6667	6667

Note: Games with a single goal difference after regular time. 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 the opponent's goal line in the stoppage time. Control variables are also interacted with the *Closed* dummy. Standard errors clustered at home team level.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

In the binary setting (Column 3), we once again find that when there is a sizable rank gap in favor of the home team, the bias is twice the size compared to a case with no difference (12.25 vs. 12.25 + 17.35 = 29.6 s). In the linear model (Column 4), for equally ranked teams the bias is 14.29 s and the slope is −0.8. So the lowest ranked team actually has a -1-s (i.e., negative) bias when playing against the top team, while the top team enjoys a 29.5-s benefit against the lowest one.

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. The results are rather stable across leagues as shown by Figure A4 in the Appendix.

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, perceived quality may be better captured by the so-called Elo rating of the teams. The Elo rating system, originating from chess, is based on the past performance of the teams. The relative rating of the two teams is designed to capture the expected outcome of the game.

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 and using values from Transfermarkt.[21]

Third, another monetary measure is revenues generated by the club. Top revenues over 10 years is an indicator of being among the 20 teams that have generated the highest average revenue over 10 years. It is a binary variable by design.

In Table 3 we reproduced key results with alternative influential team definitions. In Model (1), *Home favorite* dummy is defined as the home team being one of the 20 richest teams as per Deloitte Football Money League. In Model (2), matches where home minus away Elo rating differences are in the top quartile within the given season and league are defined as games with home favorite. Model (3) defines Home favorite games as those where the difference between the squad value of the home and away team belongs to the top quartile within the given season and league. Our findings are robust to applying any of these definitions instead of our core metric.

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 affect referee decisions. Team influence is correlated with many observable characteristics: stadium attendance and capacity or the type of stadium teams have.

TABLE 3 Robustness home-team bias heterogeneity.

	(1)	(2)	(3)
Home lose	12.35***	10.55***	11.03***
	(1.74)	(1.75)	(1.69)
Home favorite	−2.84	−7.08***	−2.68
	(4.03)	(2.20)	(2.30)
Home lose × Home favorite	10.69**	13.41***	12.62***
	(4.32)	(4.01)	(3.93)
Controls	Yes	Yes	Yes
League FE	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes
R^2	0.46	0.46	0.46
Favorite definition	20 richest	ELO diff	Squad value diff
Observations	6667	6667	6667

Note: Games with a single goal difference after regular time. 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.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Importantly, the variation in attendance and the attendance-to-capacity ratio in particular are both strongly correlated with the popularity of both home and away teams. In our data, 75% of the variation in the attendance rate, an earlier measure of the crowd effect (Garicano et al., 2005) is explained by the home team, the league, the season, and the popularity of the away team. As such, no proper inference can be made based on the variation in absolute or relative attendance (see Table A5 in the Appendix).

Similarly, for running tracks (Dohmen & Sauermann, 2016), the top 6 teams in Germany have no tracks in their stadium but some of the rest (like Hertha Berlin, or Nürnberg) do. Hence, crowd influence loses its explanatory power once confounders are taken into consideration.[22]

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 the UEFA Europa League games. Refereeing in these European games is the pinnacle of a referee's career, and the number of games they worked at is a key career success metric. Influential teams—especially those with regular presence in these competitions also having positions on UEFA boards—can (albeit rather informally) block unwanted referees and possibly promote preferred ones.[23]

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 $= 2011 - 12...2020 - 21$). On average, we observe a referee for 4.9 seasons in their national leagues working at 70 games, 33 of which have a one-goal difference after 90 min.

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 is 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 versus winning. Second, $Bias\_pred_{r,s}$ is the average deviation from the predicted stoppage time for referee r when the home team is losing versus winning, based on our model (3).

To analyze the relationship between European career and bias, we first estimate 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 competition in season s and 0 otherwise.

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

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

Alternatives include restricting the sample for the 7th national league season only for each referee (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.

Results are presented in Table 4. We find 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 in case they help teams that are likely to play in European competitions.

TABLE 4 Referees' favoritism bias and UEFA jobs.

Dependent variable Referee seasons Home teams	(1) Binary All All	(2) Binary All All	(3) Binary 7th All	(4) Count 7th All	(5) Binary All Influential
Mean home bias	0.0001		0.0010	0.0026	−0.0003
	(0.0002)		(0.0009)	(0.0032)	(0.0005)
Mean pred. home bias		0.0002			
		(0.0002)			
Domestic games (*N*)	0.0292**	0.0292**	0.0419**	0.2529***	0.0142
	(0.0081)	(0.0082)	(0.0173)	(0.0571)	(0.0110)
Controls	Yes	Yes	Yes	Yes	Yes
League FE	Yes	Yes	No	No	Yes
Season FE	Yes	Yes	No	No	Yes
R^2	0.38	0.38	0.36	0.46	0.40
Observations	553	553	78	78	225

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 versus 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).

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

4 | CONCLUSIONS

In this paper, we extended and revisited earlier evidence of favoritism under social pressure, as measured by home-team favoring bias among European football referees. Using extensive data from approximately 7000 football matches spanning 10 seasons of the top 5 European leagues, we find that referees support the home team by allowing the game to last longer if it's losing. The point estimate of this bias is 13 s, or 5% of the extra time. The magnitude is comparable to earlier studies from 20 years ago.

Our key contribution lies in understanding the source of social pressure pushing referees to make such decisions. Unlike previous assumptions, the bias is not due to crowds, as it persists even in the absence of home fans: it continued when games were played in empty stadiums during the Covid-19 pandemic.

Importantly, the bias is larger when it favors influential (top-ranked) teams and is especially pronounced when an influential team is losing at home to a minnow. The point estimate of influential team bias is 24–30 s, or 10% of the extra time.

To ensure that social pressure is the driver rather than corruption, we demonstrated that the favoritism bias from referees is uncorrelated with career benefits.

The favoritism bias is quite sizable: for influential teams, the 30-s extra represents almost 10% longer stoppage time. We can translate this into a financial benefit. A team that is losing at 90:00 equalizes or even wins in 7% of the matches, meaning there is a 0.7% chance that the home bias would change the points outcome of the game, on average by 1.3 points. With 38 games in a season, this adds up to almost 0.4 points. For a team like Manchester United, which spends on average 6.5 million USD per point, the bias corresponds to more than 2.5 million USD.[24]

Thus, instead of adherence to social pressure from 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 propose a more nuanced explanation: it may be the consequence of unconscious bias driven by a persuasive home team organization. People are often uncertain about their decisions, and this is when bias occurs: 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 go unpunished if it hurts a small team but may receive massive media coverage if a top team suffers.

This paper has shown that it is difficult to eliminate favoritism. Our findings indicate that there could be multiple sources of social pressure—in our case, influential organizations beyond the crowds. In recent decades, football organizations have focused on taming crowds and punishing aggressive behavior. However, the power of some clubs may have even increased, with a global audience, more money, and higher stakes. Understanding the multiple sources of social pressure can help explain why crowd control alone is insufficient to make referees cut favoritism.

While we were able to estimate influential team bias, there is room for future research in understanding how exactly influential organizations use their star power to exert social pressure.

ACKNOWLEDGMENTS

We thank Miklós Koren, Mats Köster, László Mátyás, Gianmarco Ottaviano, Thomas Peeters, Ádám Szeidl, two anonymous referees, and seminar participants at Erasmus Center for Applied Sports Economics, Central European University, and Corvinus University for useful comments and suggestions.

DATA AVAILABILITY STATEMENT

Data and code to reproduce exhibits in the paper are shared openly and are available at https://doi.org/10.3886/E195463V5 (Bekes et al., 2024).

ORCID

Gábor Békés https://orcid.org/0000-0002-6331-4408
Endre Borza https://orcid.org/0000-0002-8804-4520

ENDNOTES

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

[2] The discretionary power empowers referees to set stoppage time at the end of each half to compensate for time lost due to various events such as injuries. See Section A1 of the Appendix for more details.

[3] 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-s bias found by Dohmen (2008) in the German Bundesliga for the 1992/93 to 2003/04 period, and substantially smaller than the 110 s Garicano et al. (2005) found for the two Spanish seasons of 1994/95 and 1998/99.

[4] For anecdotal evidence, see www.theguardian.com/football/2015/mar/28/referees-football-match-day-routine-sport and www.as.com/en/2017/02/27/soccer/1488229755_283818.html.

[5] Conformity has a large literature in psychology, see Cialdini and Trost (1998), with conformity to stars, in particular, discussed in many contexts such as body image in Shorter et al. (2008).

[6] We thank an anonymous referee for this point.

[7] Sporting success and team wealth are strongly correlated in elite sports such as football, and there is low churning at the top.

[8] 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.

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

[10] 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 A1.5.

[11] See for example, Carmichael and Thomas (2005); Dawson and Dobson (2010), and for a review, Dohmen and Sauermann (2016).

[12] There are 20 teams in a league (18 in Germany), $10 \times 4 \times 20 \times 19 + 10 \times 1 \times 18 \times 17 = 18{,}260$. 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.

[13] See: Deloitte Football Money League Wikipedia page. The 20 teams include 6 English, 4 Italian, 4 German, 3 French, and 3 Spanish teams.

[14] We thank an anonymous referee for pointing out this discrepancy.

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

[16] This step excludes matches where the game stopped for at least 266 s (almost 3 min).

[17] 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 Section A1 in the Appendix.

[18] Our results refer to the 2011–2021 period. Note that our estimated coefficient for Germany is 16 s, close to what was measured earlier by Dohmen (2008) but a magnitude smaller than the one in Garicano et al. (2005). For the lack of available data from the period, it is difficult

to make a direct comparison, but the difference in estimates is not driven by modeling choices: a replication of their core model (in their Table 2, column 4) offers a similar estimate to our favored specification (Appendix, Table A4). It is possible that in the nineties there was less oversight of referee behavior in Spain.

[19] Another novelty during Covid-19 was having cooling breaks in the summer games of 2020. We checked robustness by taking these games out, and it had no impact.

[20] Alternatively, confounders may be partialled out of rank as well, only to result in a very similar graph.

[21] For example, in the 2018/19 season, Arsenal, an English Premier League team, is valued at 659 million euros making it the 6th most valuable team (while the team holds the 5th position in the points table). https://www.transfermarkt.com/premier-league/startseite/wettbewerb/GB1/plus/?saison_id=2018.

[22] Close to fans is defined as not having a track or a sizable gap between the pitch and fans. An example of a stadium with tracks is Hertha Berlin for a covered gap is West Ham's London Stadium.

[23] See a short discussion in A1.5 in the Appendix.

[24] Calculated based on the estimated salary costs and transfer fees collected from spotrac.com.

[25] The official *Laws of the Game* are available at www.theifab.com/laws-of-the-game-documents/.

[26] See, for example, the following quote by former referee Dermot Gallagher: "*(...)we've had this standardization that we're going to play 30 s per substitution, and for excessive goal celebrations we're to play another 30 s—so it starts to tot up, and this is why we find the three or 4 minutes we have on average at most games.*"—www.playtheadvantage.com/2014/05/27/how-stoppage-time-is-determined/.

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

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

[29] www.mirror.co.uk/sport/football/champions-league-plan-premier-league-26923957.

[30] www.forbes.com/sites/steveprice/2022/05/11/uefas-champions-league-changes-benefit-the-big-six-and-newcastle-united/?sh=44446ce57743.

[31] www.dailymail.co.uk/sport/sportsnews/article-6347551/Leaked-documents-claim-Manchester-City-hid-30m-UEFA-FFP-investigators.html.

[32] www.spiegel.de/international/world/football-leaks-doping-tests-and-real-madrid-a-1240035.html.

[33] Note that for this table, the parameters we estimate here are not directly meaningful, as the actual number of seconds spent on these events are already partialled out by the wasted time variable. This functional form decision allows for referee heuristics to be accounted for.

REFERENCES

Austen-Smith, D. & Fryer, R.G. (2005) An economic analysis of "acting White". *Quarterly Journal of Economics*, 120(2), 551–583. Available from: https://doi.org/10.1093/qje/120.2.551

Battaglini, M. (2002) Multiple referrals and multidimensional cheap talk. *Econometrica*, 70(4), 1379–1401. Available from: https://doi.org/10.1111/1468-0262.00336

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

Bekes, G., Borza, E. & Fleck, M. (2024) *ECIN replication package for "Favoritism under multiple sources of social pressure."* Ann Arbor: Inter-university Consortium for Political and Social Research [distributor]. Available from: https://doi.org/10.3886/E195463V5

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

Bursztyn, L. & Jensen, R. (2017) Social image and economic behavior in the field: identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9(1), 131–153. Available from: https://doi.org/10.1146/annurev-economics-063016-103625

Carmichael, F. & Thomas, D. (2005) Home-field effect and team performance: evidence from English Premiership Football. *Journal of Sports Economics*, 6(3), 264–281. Available from: https://doi.org/10.1177/1527002504266154

Caselli, M., Falco, P. & Mattera, G. (2022) When the stadium goes silent: how crowds affect the performance of discriminated groups. *Journal of Labor Economics*, 41(2), 431–451. Available from: https://doi.org/10.1086/719967

Cialdini, R.B. & Trost, M.R. (1998) Social influence: social norms, conformity and compliance. In Gilbert, G.L.D.T. & Fiske, S.T. (Eds.) *The handbook of social psychology*. McGraw-Hill, pp. 151–192.

Dawson, P. & Dobson, S. (2010) The influence of social pressure and nationality on individual decisions: evidence from the behaviour of referees. *Journal of Economic Psychology*, 31(2), 181–191. Available from: https://doi.org/10.1016/j.joep.2009.06.001

DellaVigna, S. & Gentzkow, M. (2010) Persuasion: empirical evidence. *Annual Review of Economics*, 2(1), 643–669. Available from: https://doi.org/10.1146/annurev.economics.102308.124309

DeMarzo, P.M., Vayanos, D. & Zwiebel, J. (2003) Persuasion bias, social influence, and unidimensional opinions. *Quarterly Journal of Economics*, 118(3), 909–968. Available from: https://doi.org/10.1162/00335530360698469

Dohmen, T. (2008) The influence of social forces: evidence from the behavior of football referees. *Economic Inquiry*, 46(3), 411–424. Available from: https://doi.org/10.1111/j.1465-7295.2007.00112.x

Dohmen, T. & Sauermann, J. (2016) Referee bias. *Journal of Economic Surveys*, 30(4), 679–695. Available from: https://doi.org/10.1111/joes.12106

Duggan, M. & Levitt, S.D. (2002) Winning isn't everything: corruption in sumo wrestling. *The American Economic Review*, 92(5), 1594–1605. Available from: https://doi.org/10.1257/000282802762024665

Erikstad, M.K. & Johansen, B.T. (2020) Referee bias in professional football: favoritism toward successful teams in potential penalty situations. *Frontiers in Sports and Active Living*, 2. Available from: https://doi.org/10.3389/fspor.2020.00019

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

Gauriot, R. & Page, L. (2019) Fooled by performance randomness: overrewarding luck. *The Review of Economics and Statistics*, 101(4), 658–666. Available from: https://doi.org/10.1162/rest_a_00783

Gentzkow, M. & Kamenica, E. (2016) Competition in persuasion. *The Review of Economic Studies*, 84(1), 300–322. Available from: https://doi.org/10.1093/restud/rdw052

Gong, H. (2022) The effect of the crowd on home bias: evidence from NBA games during the COVID-19 pandemic. *Journal of Sports Economics*, 23(7), 950–975. Available from: https://doi.org/10.1177/15270025211073337

Hodler, R. & Raschky, P.A. (2014) Regional favoritism. *Quarterly Journal of Economics*, 129(2), 995–1033. Available from: https://doi.org/10.1093/qje/qju004

Jamieson, J.P. (2010) The home field advantage in athletics: a meta-analysis. *Journal of Applied Social Psychology*, 40(7), 1819–1848. Available from: https://doi.org/10.1111/j.1559-1816.2010.00641.x

Morita, H. & Araki, S. (2022) Social pressure in football matches: an event study of 'remote matches' in Japan. *Applied Economics Letters*, 0(0), 1–4. Available from: https://doi.org/10.1080/13504851.2022.2066617

Parsons, C.A., Sulaeman, J., Yates, M.C. & Hamermesh, D.S. (2011) Strike three: discrimination, incentives, and evaluation. *The American Economic Review*, 101(4), 1410–1435. Available from: https://doi.org/10.1257/aer.101.4.1410

Prendergast, C. & Topel, R.H. (1996) Favoritism in organizations. *Journal of Political Economy*, 104(5), 958–978. Available from: https://doi.org/10.1086/262048

Reade, J.J., Schreyer, D. & Singleton, C. (2022) Eliminating supportive crowds reduces referee bias. *Economic Inquiry*, 60(3), 1416–1436. Available from: https://doi.org/10.1111/ecin.13063

Schwartzstein, J. & Sunderam, A. (2021) Using models to persuade. *The American Economic Review*, 111(1), 276–323. Available from: https://doi.org/10.1257/aer.20191074

Shorter, L., Brown, S.L., Quinton, S.J. & Hinton, L. (2008) Relationships between body-shape discrepancies with favored celebrities and disordered eating in young women. *Journal of Applied Social Psychology*, 38(5), 1364–1377. Available from: https://doi.org/10.1111/j.1559-1816.2008.00351.x

Vollaard, B. & van Ours, J.C. (2022) Bias in expert product reviews. *Journal of Economic Behavior & Organization*, 202, 105–118. Available from: https://doi.org/10.1016/j.jebo.2022.08.002

Wolaver, A. & Magee, C. (2022) Ghost games: crowds, referee bias, and home advantage in European football leagues. *Journal of Sport Behavior*, 45(3), 950–975.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Békés, G., Borza, E. & Fleck, M. (2024) Favoritism under multiple sources of social pressure. *Economic Inquiry*, 1–22. Available from: https://doi.org/10.1111/ecin.13245

APPENDIX A

A1 | Football rules and practices

A1.1 | 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

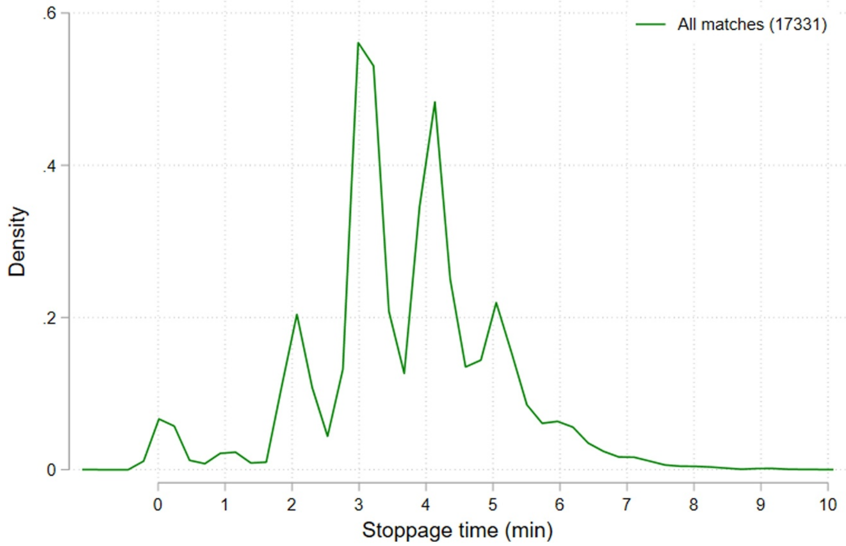


FIGURE A1 Distribution of additional time.

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.[25] Thus, the referee decision is driven by a set of detailed rules regarding which events may generate an extension. Still, the exact length is determined by the referee.[26]

Figure A1 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 min.

A1.2 | Timestamp of the game end

The event data, collected from whoscored.com, indicated a time stamp for an event type called end, at the end of both halves for each game. There may be some delay in the timestamp of this event compared to the actual time of the final whistle, as the system of whoscored.com takes a few seconds to record the final result of the game. To check if this could pose a bias, we took a random subset of 30 games from our analysis sample, games that stood at a one-goal difference before stoppage time across all leagues and seasons. We watched recordings of 30 matches: the game-end timestamp was very close to the actual final whistle, with only minor deviations (ranging from 0 to 6 s, with an average of 2 s). The example table is available in the paper's code repository.

To see any possible bias, we calculated the average difference between the final whistle and our measured "end" time across these groups, only to find them the same (2.29 s for the home lead and 2.31 s for the away lead).

A1.3 | 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, potential penalty situations, and potential 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.

A1.4 | Leagues

In a sporting season, teams play in their national leagues twice with every other team, once at home and once as a visitor. A win yields three points, a draw yields one, and a loss yields nothing. These points sum up at the end of the season to yield a final league table and create the ranking of teams. As of season 2020/2021, the first four teams from each league will play in the UEFA Champions League, and the fifth and sixth will play in the less lucrative UEFA Europa League next season. The last two or three teams will be relegated to second division. Often a single point decides about winning European spots or relegation.

A1.5 | Big team influence at UEFA

Formally, teams can have no say over referee jobs.[27] However, there were several news reports when teams tried to exert pressure on UEFA regarding referees.[28]

Another area where teams do have informal influence is the format of European games. Indeed, as early as 1997, major teams made the Champions League include runner-ups from 8 leagues, and it was later extended to include the current 4 teams of the top 5 leagues. In 2021, UEFA proposed changes in the Champions League to please the dozen most influential teams that were trying to break away and form the Super League.[29] The new Champion League format benefits the big teams (mainly English ones).[30]

Football Leaks, a blog, exposed several cases that according to UEFA regulations should have ended with certain penalties but did not in the end, such as for Manchester City and PSG.[31] There are other tax and doping cases of big teams that UEFA helped cover up.[32] These events all suggest a substantial informal influence of big teams on the decision making of UEFA.

A2 | Covid and closures

Due to the Covid-19 outbreak in Europe in the Spring of 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 was played on the 11th of March in Germany. The remaining games of the 19/20 season were played, with a more intensive schedule, starting on 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 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-19 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; and for Italy and Spain in May 2021. As Table A1 shows, this partial opening meant that the stadiums were filled

TABLE A1 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	12.6	25
France 19/20	279	1	0		
France 20/21	378	316	62	14.3	33.4
Germany 19/20	306	83	0		
Germany 20/21	306	269	37	10.8	20.9
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	11.8	20.4

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.

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 33%.

A3 | Data cleaning

For 8 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.

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, indicating maximum distance from the opponent's goal line.

A4 | Descriptive statistics

Figure A2 shows the average stoppage time by goal difference at 90:00. It shows a pattern very similar to the one documented by (Garicano et al., 2005): 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

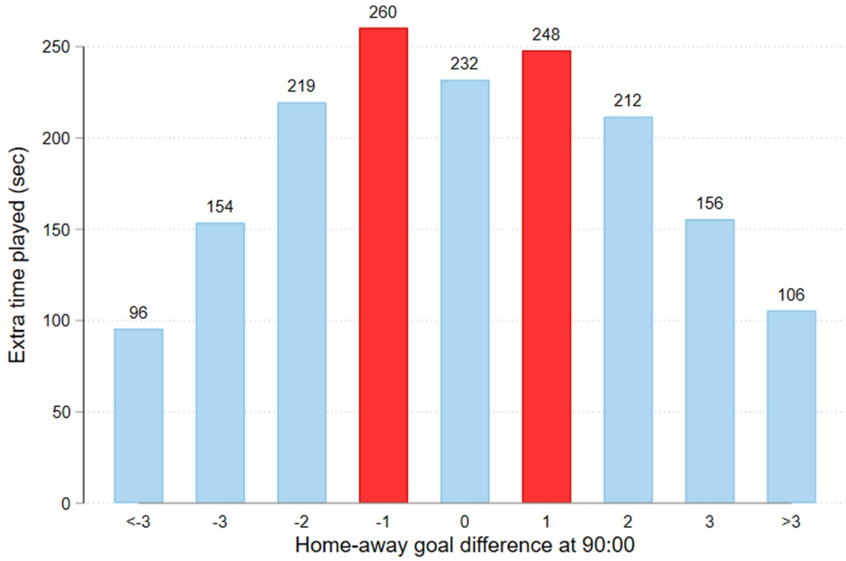


FIGURE A2 Stoppage time awarded by score margin.

TABLE A2 Summary statistics.

	Mean	SD	Median	Min	Max	*N*
Stoppage time (sec)	253.43	69.02	246	3	660	6667
Home losing (0/1)	0.43	0.50	0	0	1	6667
Wasted time (sec)	1244.42	184.55	1241	661	1964	6667
Cards	3.13	1.88	3	0	13	6667
Subs	5.74	1.14	6	2	10	6667
Fouls	14.42	4.18	14	3	33	6667
Goals	1.39	1.13	1	0	7	6667
Losing offensiveness	61.00	9.06	61.56	7.10	100.00	6661

(Continues)

TABLE A2 (Continued)

	Mean	SD	Median	Min	Max	*N*
Goals in stoppage time	0.16	0.39	0	0	3	6667
Round	19.17	10.75	19	1	38	6667
Elo ranking difference (home-away)	−8.29	145.12	−5.70	−482.57	477.84	6667
Home rank	10.66	5.55	11.00	1.00	20.00	6667
Away rank	10.23	5.62	10.00	1.00	20.00	6667
Home-away rank difference	0.43	7.93	1	−19	19	6667
Home team value (m EUR)	8.28	9.44	5	1	64	6667
Away team value (m EUR)	9.03	10.21	5	1	64	6667
Value difference (H-A)	−0.75	12.63	−0.17	−61.60	61.60	6667
VAR (0/1)	0.30	0.46	0.00	0.00	1.00	6667

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. See Section 2.2 for a detailed description of the variables.

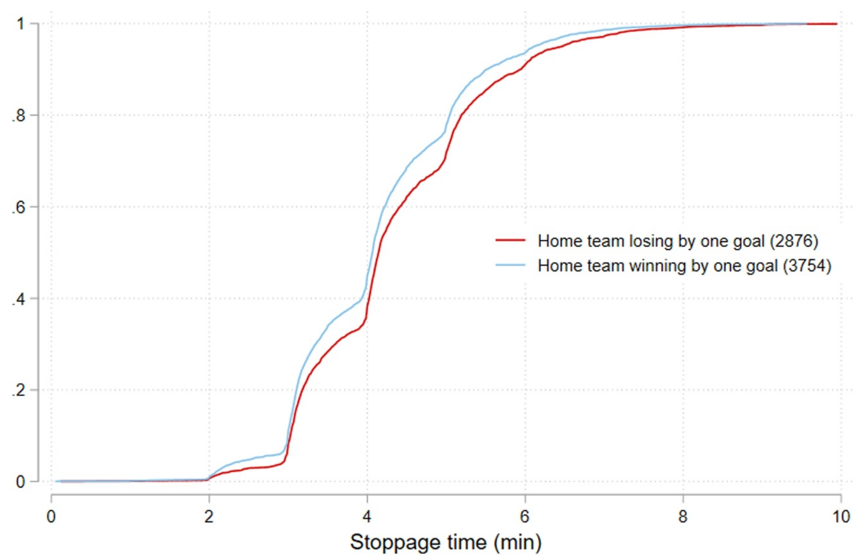


FIGURE A3 Comparing the cumulative distribution of home team leading versus losing.

already settled, and in this case, referees tend to blow the final whistle significantly earlier, serving the interest of both teams.

Table A2 presents some informative descriptive statistics about the variables included in the analysis.

Figure A3 looks at densities only for games where the home team is winning or losing by a single goal, and we can see that games won by the home team are characterized by shorter stoppage time than those lost by the home team. A Kolmogorov-Smirnov test also confirms the statistical significance of the difference between the two distributions.

TABLE A3 Detailed regressions indicating home-team bias.

	(1)	(2)	(3)	(4)	(5)
Home lose	12.40***	13.38***	11.65***	13.41***	13.40***
	(2.07)	(1.91)	(1.65)	(1.65)	(1.62)
Wasted time		0.18***	0.17***	0.18***	0.18***
		(0.01)	(0.01)	(0.01)	(0.01)
Cards		3.42***	4.75***	4.40***	4.28***
		(0.58)	(0.48)	(0.46)	(0.45)
Subs		4.69***	7.95***	7.39***	7.94***
		(0.85)	(0.84)	(0.81)	(0.78)
Fouls		−3.01***	−2.03***	−2.07***	−1.84***
		(0.28)	(0.21)	(0.21)	(0.21)
Goals		−4.21***	−4.12***	−4.53***	−4.18***
		(0.74)	(0.70)	(0.71)	(0.74)
Losing offensiveness		0.18**	0.25***	0.29***	0.26***
		(0.08)	(0.07)	(0.07)	(0.07)
ET goals		19.13***	18.41***	18.21***	18.04***
		(2.18)	(2.06)	(2.05)	(1.93)
VAR		13.66***	14.46***	14.60***	18.05***
		(2.30)	(1.93)	(1.94)	(2.16)
Round		−0.15**	−0.21***	−0.21***	−0.21***
		(0.06)	(0.06)	(0.06)	(0.06)
Constant	248.04***	13.93			
	(2.48)	(9.21)			
League FE	No	No	Yes	Yes	Yes
Home team FE	No	No	No	Yes	Yes
Referee FE	No	No	No	No	Yes
Observations	6667	6667	6667	6667	6646
R^2	0.01	0.33	0.42	0.45	0.49

Note: Standard errors clustered at home team level.

A5 | Additional tables

First, in Table A3, we present the detailed regression results with all control variables.[33] Then, we show a replication of the core table in (Garicano et al., 2005) followed by a table on attendance in Tables A4 and A5, respectively.

TABLE A4 Replication of home bias from Garicano et al. (2005).

	(1)	(2)	(3)	(4)	(5)
Home lose	9.04***	8.94***	8.99***	3.22	11.18*
	(1.69)	(1.75)	(2.97)	(3.34)	(6.20)
Yellow	7.61***	7.71***	7.67***	8.08***	8.06***
	(0.46)	(0.48)	(0.48)	(0.53)	(0.52)
Red	18.38***	18.95***	18.87***		
	(1.75)	(1.67)	(1.73)		
Subs	14.21***	13.92***	14.45***		
	(1.08)	(1.10)	(1.10)		
Home value million	−0.65**	−0.63**	−0.49***	−0.83***	−0.80***
	(0.27)	(0.27)	(0.13)	(0.17)	(0.17)
Away value million	−0.46***	−0.45***	−0.45***	−0.57***	−0.61***
	(0.10)	(0.11)	(0.10)	(0.10)	(0.10)
Home rank	−0.15	−0.19	−0.45*	−0.48*	−0.47*
	(0.32)	(0.30)	(0.25)	(0.28)	(0.28)
Home-away rank diff	0.28*	0.25	0.23	0.34**	0.32*
	(0.16)	(0.16)	(0.16)	(0.17)	(0.17)
Round			−0.16**	−0.21***	−0.22***
			(0.08)	(0.08)	(0.08)
Home lose × Round			0.00		
			(0.12)		
Attendance 1000				234.66***	141.98
				(88.76)	(98.15)
Home lose × Attendance 1000				205.56*	301.01**
				(108.97)	(120.10)
Attendance/Capacity (%)					0.19**
					(0.07)
Home lose × Attendance/Capacity (%)					−0.14
					(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	0.31	0.36	0.29	0.26	0.26
Model	T2C4	T2C6	T5C4	T6C3	T6C4
Observations	5876	5853	5876	5390	5390

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 A5 Explaining attendance/capacity ratio.

	(1)	(2)	(3)
Away top 6	6.09***		
	(0.47)		
Home-away rank diff ≤ −10		−1.68***	
		(0.48)	
Home-away rank diff			0.28***
			(0.04)
Season FE	Yes	Yes	Yes
League FE	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes
R^2	0.76	0.75	0.76
Observations	16,075	16,075	16,075

Note: Standard errors clustered at home team level.

A6 | 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 1 and Model (1) of Table 2 for each league separately (without league fixed effects). The estimated coefficients are presented in Tables A6 and A7.

We find that all main results are highly robust across leagues. First, the home-team bias is very similar, ranging between 9.4 and 17.9 s (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 −5.9 and + 6.5 s, 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 6.8 and 16.4 s (neither is statistically different). Note, however, that broken down by leagues, this result lacks statistical power, and the results are not always significant. This is because top teams rarely lose at home, and the number of observations by league is too small. Figure A4 summarizes our main findings.

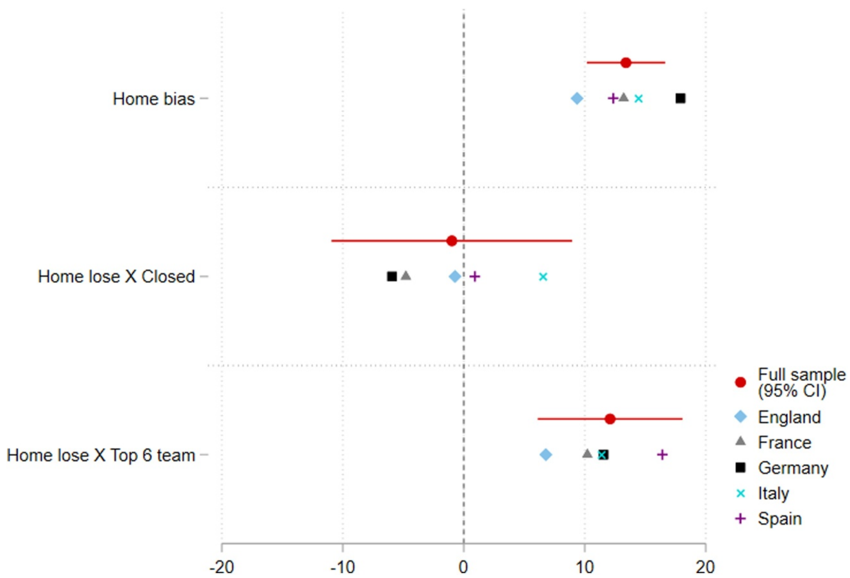


FIGURE A4 Estimated regression coefficients by league.

Tables A6 and A7 show the regressions visualized on Figure A4.

TABLE A6 Regressions by league.

	(1)	(2)	(3)	(4)	(5)
Home lose	9.37***	13.22***	17.91***	14.43***	12.36***
	(3.22)	(2.62)	(5.33)	(2.82)	(3.64)
Home lose × Closed	−0.72	−4.79	−5.93	6.54	0.92
	(9.81)	(10.92)	(11.55)	(7.99)	(12.93)
Controls	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes	Yes
Observations	1370	1397	1055	1432	1413
R^2	0.37	0.44	0.43	0.40	0.41
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 A7 Regressions by league.

	(1)	(2)	(3)	(4)	(5)
Home lose	7.74*	10.80***	15.12***	12.68***	8.88**
	(4.13)	(3.15)	(4.89)	(3.05)	(4.27)
Home lose × Home top 6	6.79	10.21*	11.55	11.38*	16.42**
	(6.89)	(5.77)	(7.89)	(5.67)	(6.26)
Controls	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes	Yes	Yes
R^2	0.37	0.44	0.43	0.40	0.41
League	England	France	Germany	Italy	Spain
Observations	1370	1397	1055	1432	1413

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.