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] *Centre for Economic Policy Research, London, United Kingdom*

Draft, v1.02 – 05 May 2022. See latest version HERE.

Abstract

This paper investigates potential mechanisms of social pressure leading to biased decisions favoring one of the parties despite detailed rules guiding them. We revisit and extend Garicano, Palacios-Huerta, and Prendergast (2005) who found that football referees make biased decisions in setting the length of stoppage time, in favor of the home team. In their argument, the source of social pressure is to satisfy the crowd in the stadium. Using exogenous variation in crowd size due to closed games prompted by the Covid-19 outbreak, along with a substantially finer and wider dataset, we revisit the setup and propose a new mechanism. We first confirm that the home team bias exists, albeit smaller in magnitude, despite all games being televised and data publicly shared. Importantly, we show that this bias is uncorrelated with home crowd size including events with empty stadia. Instead, we find that the bias mostly comes from favoring big teams of the leagues. Home bias is estimated to be over twice as much for the top teams, and almost zero for the weakest ones. This suggests that referees remain to be influenced by social pressure, but instead of crowds, conformity to and influence by host organizations are linked with this bias. Nevertheless we find no evidence of bias behaviour leading to better career outcomes in Europe.

Keywords: social pressure, favoritism, football

JEL-codes: D71, Z20, C21

1 Introduction

Favoritism, the practice of giving unfair preferential treatment to one entity at the expense of another, has broad consequences in a wide range of situations. In this paper we are concerned with how decision makers, such as judges or referees, make biased decisions that are not warranted by the rightful determinants of the decisions, and that systematically help one party to the detriment of another party.

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

We are particularly interested in the source of this favoritism. In some cases, the bias may be attributed to the preferences of the decision maker. This covers in-group bias, the tendency to prefer entities with the same gender, nationality, language, or race, among other things. This also includes corruption, when the decision is motivated by financial or other *quid pro quo* rewards. In other cases, however, the bias may be attributed to social pressure in the sense that the individual decision is influenced by the surrounding group. This influence can take various forms[1], but their common attribute is that some expectations are communicated towards the decision maker, with a shared understanding of gains to conformity.

This paper relates to the literature in several ways. Bursztyn and Jensen (2017) provide a review on how social image and social pressure affect individual behavior, as well as on the mechanism and forces behind this pressure in different areas such as education, consumption, or effort in workplace.

The different forms of social forces impacting individual behavior have deep theoretical foundations in the economics literature (Akerlof, 1980; Manski, 2000; Bénabou and Tirole, 2006). In particular, classic social psychology experiment results of conformity (Asch, 1951) have been incorporated into economic theory by Bernheim (1994), showing that individuals are willing to suppress their preferences and conform to the social pressure in order to avoid popularity impairment through departure from the group norms.

Although the importance of social pressure has been already shown in seminal laboratory experiments (Asch, 1951; Milgram, 1963) and field experiments (Gerber, Green, and Larimer, 2008; DellaVigna, List, and Malmendier, 2012), it is difficult to distinguish social motives from persuasion or rational diffusion of information when it comes to observational data (DellaVigna, 2009). In our analysis, a unique setting in which social pressure motives can be identified, combined with a novel, high frequency data and with a natural experiment, we can identify social pressure, and disentangle the mechanisms behind this pressure.

It is typically hard to measure and examine the nature of social pressure and its link to favoritism. In this paper, we combine a unique setting in which social pressure motives can be identified with novel, high frequency data and with a natural experiment that provides exogenous variation in social pressure.

Our exercise is based on the rule in football that gives referees the discretion over the amount of stoppage time allowance at the end of the game. We build on Garicano et al. (2005) who established that referees make biased decisions in the length of stoppage time in favor of the home team, called the home-team bias. In particular, they tend to add more extra minutes when the home team is losing by one goal compared to the case when the home team is winning by one goal, thus increasing the chance for them to achieve a favorable outcome through a goal scored in stoppage time. Using data from two seasons (1994/1995 and 1998/1999) of the Spanish *La Liga*, they argue that the motive through which the bias operates is to satisfy the crowd in the stadium.

[1] See the definition of social pressure according to the American Psychological Association: "*Social pressure is the exertion of influence on a person or group by another person or group. Like group pressure, social pressure includes rational argument and persuasion (informational influence), calls for conformity (normative influence), and direct forms of influence, such as demands, threats, or personal attacks on the one hand and promises of rewards or social approval on the other (interpersonal influence)*" (VandenBos, 2007).

This position found some support by analyzing the size and composition of attendance (Dohmen, 2008), the distance of attendees to the field (Scoppa, 2008; Buraimo, Forrest, and Simmons, 2010; Dawson and Dobson, 2010) or by experimentally manipulating crowd noise (Nevill, Balmer, and Williams, 2002).

To better understand the nature of social pressure, we combine three innovations. First, instead of variation in attendance, we use stadium closures as natural experiment to identify the role of crowds. Second, with high frequency data, we can better estimate the home-team bias, its sources and its external validity. Third, with data covering 10 seasons in 5 countries, we get enough variation to estimate a novel mechanism related to the power of teams. Furthermore, we are able to condition on other biases by referees and also reject a quid-pro-quo explanation.

In this paper, we use a novel, fine granularity dataset on 10 seasons of the 5 top-flight leagues in Europe: the English *Premier League*, the Spanish *La Liga*, the Italian *Serie A*, the German *Bundesliga 1*, and the French *Ligue 1*. Our data is an event-by-event data that records each action (such as passes, penalties, injuries) along with a time stamp of that event.

The stoppage time compensates for the time lost for a variety of events (see Section 5.1 of the Appendix). Although referees influence the outcome through a very large number of decisions during the game (such as fouls, disciplinary sanctions such as yellow and red cards, penalties), these decisions are always intertwined with the actions of the players in each situation, thus it is difficult to empirically disentangle referee behavior from player behavior in these cases[2].

The length of stoppage time allowed by the referee, however, does provide us with a measurable, isolated variable of referee behavior, which allows us to draw conclusions on biased referee decisions.

The data allows us to precisely estimate favoritism in the form a home-team bias in referee decisions in football. We find that referees add 15 seconds more to stoppage time when the home team is losing by one goal compared to when the away team losing by one goal. This suggests favoritism towards home team, as more stoppage time allows more chances for the losing side to catch up. This difference is robust to including a great deal of control variables such an estimate of wasted time as well as league, home team, and referee fixed effects.

This finding offers external validity to the original evidence for referee favoritism established by Garicano et al. (2005), despite the fact that matches have become much more transparent since the period they analyzed, due to extensive media coverage and high-tech analytical solutions, and referee compensation changed to offer more incentive for a professional refereeing bi

Furthermore, it may be argued that change in regulations and a more professional game with better aligned financial incentives would have eliminated biased referee decisions as argued by Rickman and Witt (2008). Incentives were considered to drive a selection of better, less partial referees into the League. We contribute to this argument by suggesting that despite incentives a considerable home-team bias still remains in place.

We also investigate the mechanism of social pressure in favoritism. One commonly favored explanation is society at large manifested via open and public verbal expression of support. This

[2](See e.g. (Carmichael and Thomas, 2005), and Dohmen and Sauermann (2016) for an overview

may be in the form of manifestation, public discourse, or even, boycotts [3]. In the present context, the social pressure may come from the fans present in the stadium, loudly voicing opinions and support. Referees, it is argued, bow to this pressure by helping the home team. Garicano et al. (2005) found evidence supporting this position, identified using variation in relative and absolute attendance driven by predicted share of visiting fans. However, this relationship may be confounded by several factors such as team quality and popularity, or the stake of the particular game. Instead of endogenous attendance, we use a natural experiment to test if crowd size is indeed the mechanism causing favoritism (home-team bias). Exogenous variation in crowd size comes from games being played in closed stadiums following Covid-19 protocol over the course of 2020 and 2021[4]. We show that the 15 second bias remains completely unchanged even when there are no fans present. In this sense, we find evidence against crowd pressure as a source of our measure of referee home-team bias. Importantly, the pooling of data for all five leagues in our sample is necessary for this result to ensure power for hypothesis testing as the number of closed games are limited.

If not crowds, what mechanism could generate the home-team bias? We argue that the social pressure could emanate from the professional organization of the team itself. This is another indirect social pressure that may come from the home team staff and infrastructure surrounding the referee around the match, but can also take the form of implicit expectations coming from the fans of the club around the world, as well as the representatives of the club. Consistent with the latter channel, we find that home-team bias turns out to be magnified when it helps one of the top teams in the league, and attenuated when it would help the opponents of these top teams. Thus, referees tend to support groups with the largest financial and informal clout. This suggests that favoritism is present in a more traditional form: in conformity to the powerful host.[5].

Finally we discuss the possible sources of home team support bias for more powerful organizations. Our data helps us tell apart a possible quid pro quo argument: we find no evidence that referees help powerful teams in hope of their future career advancement. We also find that more experienced referees have smaller bias.

Combined together, these results suggest that referees yield to pressure by powerful organizations even if they don't expect financial gains. This observation is related to a stream of work in social psychology: people internalize expectations and are affected by power[6]. Our setting is very special: it is an extremely open and transparent situation. Thus, in more opaque settings, we expect judges to yield to power even without corruption.

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 bias, show that is not driven by crowd in the stadium, speculate about possible channels and argue that

[3] On boycotts, see (Klein, Smith, and John, 2004) [MORE]

[4] Using other outcomes, such as disciplinary actions, several studies suggested that the game changed during covid (Bryson, Dolton, Reade, Schreyer, and Singleton, 2021; Scoppa, 2021; Reade, Schreyer, and Singleton, *forthcoming*).

[5] The idea of big 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.

[6] PSYCH LIT HERE

an important factor is favoritism to influential teams. Finally, we also show the stability of our results over time and across countries. Section 4 concludes.

2 Data and model

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.

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 all leagues but Germany, there are 20 teams, 18 in Germany. Each team hosts every other team once in a season in its own stadium. This coverage would imply 18,260 games[7]. Our actual sample is slightly smaller. First, due to the Covid-19 outbreak, season 2019/20 of French Ligue 1 finished prematurely, with only 279 out of 380 games played. Due to some deficiencies in our main data sources (see below), we lost another 41 games. Thus, our final sample consists of 18,118 games.

Our dataset was created using several data sources. First, we use an an event level description of every game scraped from online sources. Each event has a type and a timestamp (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 tracked every third or fourth second on average.

We aggregated this event data to game-half level by taking the sum of events flagged as goal, foul, substitution, yellow or red card, separately for each half of the game. To take into account the events taking place during stoppage time (in particular to be able to control for goals during stoppage time, as a crucial event that changes the outcome of the game), we do this aggregation separately for regular time and stoppage time.

The second set of data is at the game level, collected from whoscored.com. It includes the venue of the game, attendance in the stadium, goals by home and away teams, referee ID, date, and time.

The third set of data concerns creating measures of favorite teams to win games. For this dataset, we combined three additional sources: betting odds, financial revenues, and squad values.

Information on pre-match odds come from of the betting site bet365.com for home win, draw, and away win[8] Next, we flagged the top 20 teams from the 5 leagues generating the most revenue according to the average yearly ranking in the Deloitte Football Money League in the past 10 seasons. The 20 teams include 6 English, 4 Italian, 4 German, 3 French, and 3 Spanish

[7] $10 \times 4 \times 20 \times 19 \; + 10 \times 1 \times 18 \times 17 = 18260$

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

teams (for details, see the Appendix). Finally, we collected data to estimate the squad value of each team for every season, using historical player valuations from transfermarkt.com.

2.2 Model and variables

In our empirical strategy we compare the stoppage time in games where the home team is winning by the smallest margin of one goal to that in games where the away team is winning by one goal. Thus, our analysis includes the matches where the absolute goal difference at the end of regular time (90:00) is exactly one goal. Among these matches ($N = 7021$ out of 18,118), we flag games where away team is winning before the stoppage time with a *Home losing* indicator variable. Our main outcome variable of interest, the length of stoppage time, is directly accessible from the event data based on the timestamp of the event indicating the end of the game, that is the minute and second of the referee's final whistle.

To be able to make a causal claim, we need to exclude confounding factors as well as mechanisms of reverse causality. Most importantly, 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.

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 ball out of play, referees rather tend to use heuristics such as the number of these key events when they make their decisions on the extra time. 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.

There is a technological change during our period that may also affect the mechanisms underlying the bias of our interest. The Video Assistant Referee (VAR) system, one of the important recent innovations in football, has been introduced 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. 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.[9]

[9] VAR has been in operation since season 2017/18 in Germany and Italy, since season 2018/19 in France and

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 more likely to attack, and the away team is more likely to lose. 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 of these passes from the attacking goal line. This variable (called *Losing offensiveness*) is measured in units of distance from the attacking goal line, on a scale from 0 to 100.[10]

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 is, the more stake there is in the game. It is reasonable to assume that the pattern of the matches changes as season advances – by controlling for this factor, we rule out this potential confounder.

Table 1 describes the variables defined above. 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 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). Our estimated wasted time in the second half varies between 9 minutes and more than 33 minutes – the latter corresponding to a severe head injury of the goalkeeper, which involves on-field medical care and, thus, a long period of play stop. The average wasted time is 19 minutes (out of 45+4), which is equivalent with about 60% effective playing time, in line with the general findings of sports press.[11]

Table 1: 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
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 only events happening in the second half of the games (where this is relevant). See Section 2.2 for a detailed description of the variables.

Spain, and since season 2019/20 in England.

[10] 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.

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

We also include different sets of fixed effects in our models. To capture general playing style and other differences, we include both league and home team fixed effects. Differences in the extra time length customs of referees as well as their attitude towards home and away teams can be captured by referee fixed effects. It cannot be excluded, however, that allocation of referees is not random, and certain types of games are given to referees with habits of longer or shorter stoppage time. This potential mechanism is especially important for the interpretation of our results regarding bias towards influential teams. Thus, although it does not significantly affect any of our results, our preferred specifications do not include referee fixed effects.

3 Results

In this section, we present results from our step-by-step empirical investigation. First, we confirm a referee bias supporting the home team. Second, we show that it is not driven by home crowds. Third we ask what could be the key mediator variable and suggest that heterogeneity in team influence to be a key one. Finally we show stability of results across leagues.

3.1 Measure of the home-team bias

To establish the basic facts, we illustrate that the length of stoppage time does depend on the result in a systematic way. This analysis, originated from Garicano et al. (2005) has been replicated by multiple studies from various leagues, with mixed results (Dohmen and Sauermann, 2016) – to our knowledge, our analysis is the first one to cover a massive number of leagues and seasons.

The basic proposition is illustrated in Figure 1, where we plot the average stoppage time by goal difference at the end of the regular time.

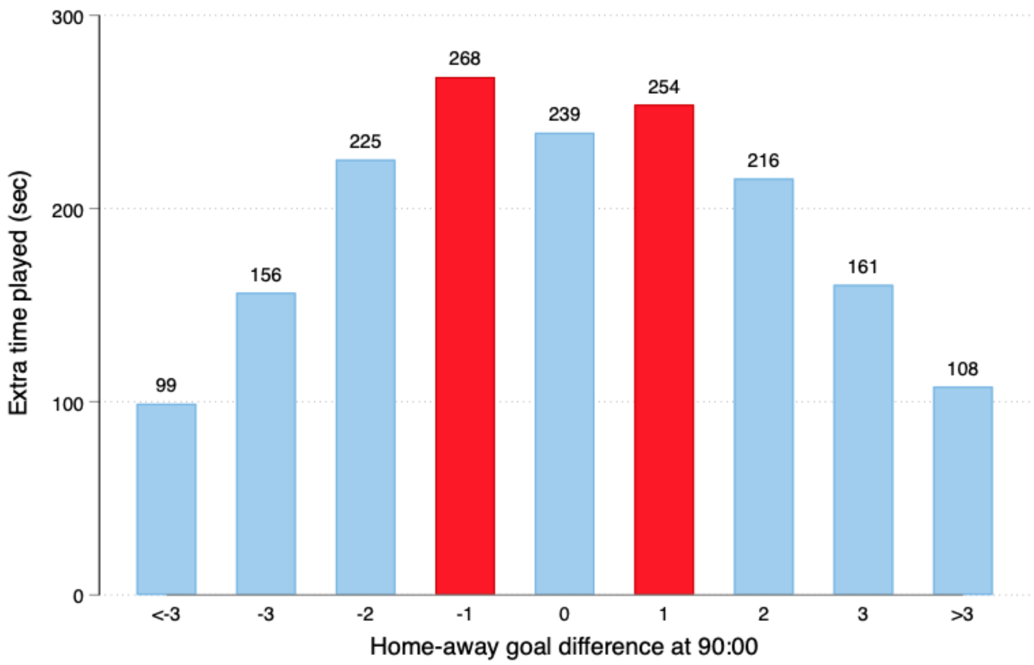
Referees do seem to take into account the result when deciding on stoppage time. It is longer the tighter the match is: the average stoppage time for one goal difference is 260 seconds, and it is 219, 159, and 105 seconds for a two, three, or larger goal difference, respectively. Since goals in stoppage 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. A draw may be useful for one or the other team depending on many aspects of the game and the season, so we do not base our analysis on drawing games.

The cleanest comparison to study referee bias comes form comparing a one-goal lead by either team. Thus, the pattern of interest for us is the 14 seconds gap between the two highlighted columns: the additional time seems to be extended in case the home team is losing compared to the case when the home team is winning by a small margin. For larger margins we see no difference or a difference for the opposite direction. For more details about stoppage time, see Section 5.1 in the Appendix. In what follows, we explore the causes and mechanisms behind this difference.

As a baseline model, we estimate the following model with OLS[12]:

[12] We cluster our standard errors at the home team level for all models. This is a more conservative alternative

Figure 1: Average stoppage time by goal difference at 90:00



$$Extra\ time_{h,a,s} = \beta Home\ lose_{h,a,s} + \gamma Controls_{h,a,s} + \delta_{league} + \eta_{hometeam_{h,s}} (+\theta_{referee_{h,a,s}}) + \epsilon_{h,a,s}, \tag{1}$$

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 teams do not move across league, home and away team ID also identifies the league.) In a season, each team plays with every other both at its home stadium and away. Thus, within a season, the h, a dyad uniquely identifies a game.

We run all our match-level regressions including only matches with a 1-goal difference at the end of regular time (90:00). That is, we compare games where the home team is losing by one goal with games where the home team is winning by one goal at the time when the referee makes its decision on the amount of indicated stoppage time.

We use a rich set of control variables. Along with our wasted time measure (the most precise estimation determining the justifiable length of stoppage time), we also include the number of goals, substitutions, cards, and fouls in the second half – factors that influence referees' stoppage time decision as rules of thumb. We separately control for the number of goals in extra time, to take into account the possibility that this can change the structure of the extra time. We also control for *Losing offensiveness*, and add a binary variable that captures whether VAR was in operation in the given match. In some specifications, we include three sets of fixed effects: league fixed effects δ_{league}, home team fixed effects $\eta_{hometeam}$, and referee fixed effects $\theta_{referee}$.

We find that referees tend to help home team by letting the game played longer by around

to a team–season two-way clustering.

13 seconds, on average. Table 2 reports the result of several model specifications, each using a different set of control variables and fixed effects. Each of these estimations finds a robust, 11.1 to 14.3 second bias.

Table 2: 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.

We believe that $Controls_{h,a,s}$ variables[13] (events, wasted time, VAR, losing offensiveness), home team and league dummies δ_{league} and $\eta_{hometeam}$ are confounding variables, while referee fixed effects $\theta_{referee}$ may be part of the mechanism through which bias works. Thus our preferred specification is Model (4). As Table 2 suggests, however, these decisions do not affect our conclusion. In our subsequent analysis, we specify our models without referee fixed effects, though all our results are robust to including them.

3.2 The home-team bias and crowd

Garicano et al. (2005) argued that the supporting crowd at the stadium makes referees feel a pressure to help the home team. Identification came from comparing attendance to capacity rates across games, arguing that higher attendance must include visitors and hence, a lower pressure. Others in the literature also assume that the source of the home-team bias is the supporters cheering for the home team.

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 Covid-19 pandemic as an exogenous variation in stadium attendance to assess whether the bias can be attributed to the fans present in the stadium.

We believe that this approach is superior to controlling for the size and composition of

[13] See Table 7 in Appendix for how these variables affect stoppage time.

supporters in the stadium.[14] 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.

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 leagues continued around May-June, with matches played behind closed doors. In Spring 2021, partial opening allowed to fill the stadiums to 10-30% of capacity. We created two indicator variables, $Closed = 1$ when attendance is zero, and $Covid = 1$ that also include very low attendance games in 2021. As Table 3 shows, though, the difference between these two categories is minor, since only a very limited number of matches has been played with attendance since the Covid outbreak in our sample.

Table 3: 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.

The natural experiment, induced by the outbreak of the Covid-19 pandemic, allows us to test the hypothesis, articulated by Garicano et al. (2005), among others, 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. In particular, in Table 4 we estimate

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

where our treatment indicator $T_{h,a,s}$ flags only closed matches in Models (2) and (3), and all Covid matches (including those with restricted number of fans) in Models (4) and (5). In both models, all control variables are also interacted with the *Covid* or *Closed* dummy to capture that referrees may take different amount of time when dealing with rule based decisions.

[14]Note that home team, league, and season, along with an indicator of a top away team explains more than 75% of the variation in the attendance rate across all matches in all sample; see Table 12 in Appendix.

We find 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 at present day, while referees remain biased in supporting the home team when losing, their actions cannot be driven by crowd size. Even in the super extreme and exogenously driven case of closed games, the home-team bias remains as it was before. Thus, crowds cannot be the main driver. (We do not claim it never was. Maybe in the past (including the period examined by Garicano et al. (2005)) referees were indeed affected but more oversight regarding crowds helped[15].)

Table 4: Regressions indicating no crowd effect

	(1) Extra secs	(2) Extra secs	(3) Extra secs	(4) Extra secs	(5) Extra secs
Home lose	13.10***	11.60***	13.22***	11.57***	13.17***
	(1.77)	(1.76)	(1.77)	(1.77)	(1.77)
Home lose × Covid		-1.94	-1.35		
		(5.31)	(5.46)		
Home lose × Closed				-1.68	-0.71
				(5.22)	(5.44)
Controls	Yes	Yes	Yes	Yes	Yes
League FE	Yes	Yes	Yes	Yes	Yes
Home team FE	Yes	No	Yes	No	Yes
R^2	0.42	0.39	0.41	0.39	0.41
Period	pre-Covid				
Observations	6151	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 extra time. All these variables are also interacted with the *Covid* or *Closed* dummy.

This is an unexpected result given the accepted wisdom that social pressure stems from crowds. It is, however, very robust to specifications and also across leagues – see Figure 3 in Section 3.5 below.

3.3 Sources of home-team bias without crowds

This zero crowd effect result has striking consequences, as it suggests that the pressure on the referee to help the home team is not a consequence of the home supporters in the stadium. Even without fans cheering in the stadium, the referees seem to internalize the preferences of the surrounding environment. Several factors may explain this, and we may only speculate about

[15] Note that our estimated coefficient is 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 not country specific, we found it across all leagues. Thus, it is most likely time: in the nineties, check on referees was indeed weaker.

the exact nature.

Hospitality – On match day, referees arrive at the stadium early before kickoff, exposing themselves to a number of administrative staff associated with the home team, the local press, and possibly even fans around the stadium. According to some reports, referees tend to get to the town one day before the match, and spend a night in a local hotel.[16] This fact increases the exposure to local ambience even more. Moreover, 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. It is also not uncommon that upon departure, they receive a formal gift consisting of merchandise products of the home team.[17]

Attention – It may be about avoiding a backlash: as the team as well as supporters pay greater attention to home games (also expect better outcome on average, due to the home field advantage in general), any error from the referee may induce great outrage among members of the host organization, so referees would err on the home team side.

All these factors point towards the existence of social pressure mechanisms stemming from the team and the sport organization that may explain the bias pattern even in absence of crowds[18].

If this is indeed the case, and influence comes from the organizations, we shall expect teams with a greater power to have a relatively larger influence. This would manifest in a home-team bias that is function of team influence. This is the topic for the next section.

3.4 The home-team bias and 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 wealth and quality. Wealth could be measured as actual financial prowess, the value of current players as well as wealth in 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 because of its simplicity. 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).[19] Nevertheless, all these above measures are highly correlated with a correlation coefficient between 57% and 92%.

Influence also comes from the exposure of teams that will play in UEFA European championships. From these leagues, the top 6 teams will play in either Champions League or Europa

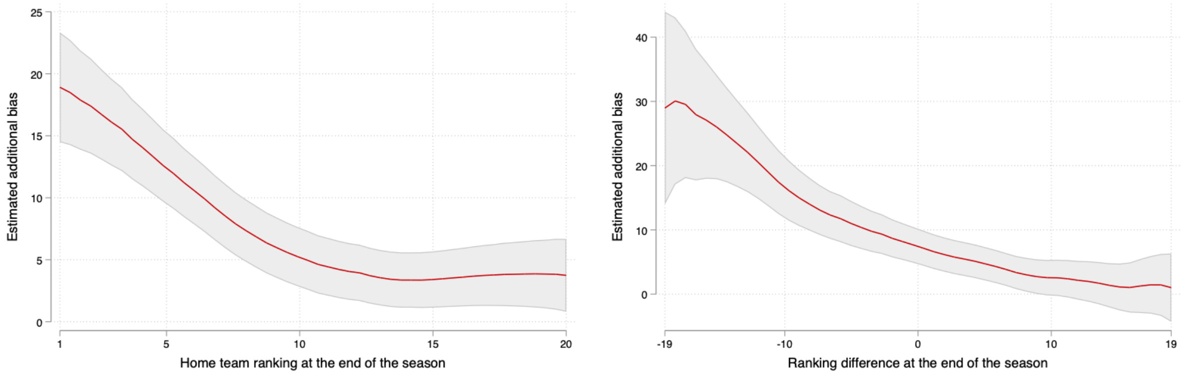
[16] See some anecdotal evidence in this article: https://www.theguardian.com/football/2015/mar/28/referees-football-match-day-routine-sport.

[17] See this article for a documentation of this phenomenon: https://en.as.com/en/2017/02/27/soccer/1488229755_283818.html.

[18] Note that there may be several other types of social pressures on referees such as traditional and social media. However, that would likely not come through the home-team bias, but instead selected teams being favored regardless of playing home or away.

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

Figure 2: Heterogeneity of home-team bias by team rankings



(a) Home team ranking

(b) Difference between home and away team ranking

Note: Local polynomial smoothing and a 95% confidence interval. Predicted extra time is the residual from a regression of extra 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.

League.

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

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

In the second step, using local polynomial smoothing regressions, we plot this bias against heterogeneity by the ranking of the home team. In Figure 2a, we see fairly strong pattern with better teams enjoying a greater home-team bias. This is especially true for the top few teams. In Figure 2b, we see a similar pattern with the difference plotted against bias, when top team play lower ranked ones, the gap is 30 seconds, but it goes down to 0 when minnows play at home versus big teams.[20]

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 per league).

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

An alternative model uses the difference between home and away team, with β_3 measur-

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

ing 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).

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

The results of both models are presented in Table 5. The simplest setup is considering the binary influential team indicator in column (1) of Table 5. Influential teams enjoy a home-team bias that is more than twice the one for non-influential ones.

Looking 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 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. 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 spots, 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 (successful) teams, especially when they play less influential (successful) ones.

To summarize, we find that influential (successful) teams benefit substantially more from referee decisions. The result is stable over time, and if anything, it is further increased in closed games. The top influential teams seems to create an environment when playing at home that makes referees more inclined to help. It is power of the host team organization and not the crowd size that determines behavior.

Table 5: Regressions indicating home-team bias heterogeneity

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

Note: 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 *Closed* dummy.

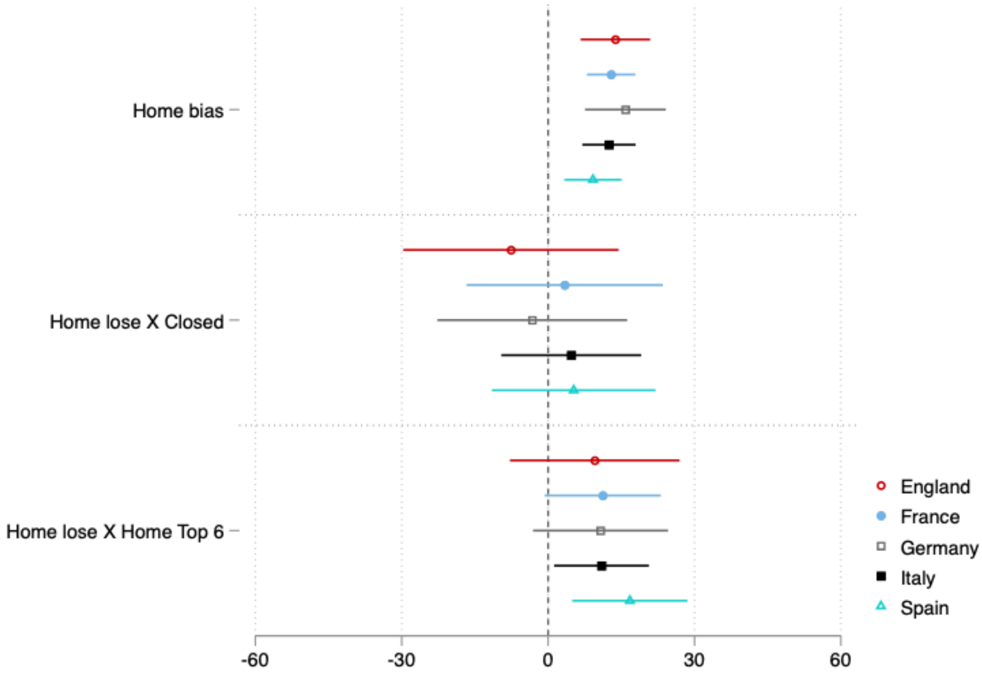
3.5 Stability across leagues

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 4 and Model (1) of Table 5 for each league separately (without league fixed effects). The estimated coefficients are presented in the Appendix Tables 9, and 10.

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 big teams rarely lose at home, and the number of observations by league is too low. Figure 3

Figure 3: Estimated regression coefficients by league



summarizes the main findings.

3.6 Stability over time and during closed games

Finally, we show that patterns observed in the pooled data, are also stable over the years. We cut the sample into 3 parts, early years (2012-2015), later years (2016-2019) and covid. (2020-21) and equation **??** with the binary indicator for influential teams (top 6 in each league) for each period.

As indicated in Table 6, before Covid, we see little difference, point estimates are higher for the second period (early period: home team losing bias is 7.5, interaction: 8.3 seconds, in later one, the home-team losing bias is 12 seconds and interaction is 9.5 seconds).

This suggests that are baseline proposition of influential teams is very stable overtime.

However, for the last years with most games played in a closed stadium we do see a different image: influential teams are given 32 seconds more when in need. The point estimate for this triple interaction term is large and points to a situation where host organizations play a very important role (covid protection), and seems to have an even greater influence on referees.

The finding that during closed games, the home-team bias gap between influential and other teams have almost tripled suggests once again social pressure from organizations and not the crowds.

Table 6: Variation across time periods

	(1)	(2)	(3)
	Extra secs	Extra secs	Extra secs
	2011/12–14/15	2015/16 – 18/19	2019/20 – 20/21
Home lose	7.460***	12.89***	7.711
	(2.574)	(2.601)	(4.943)
Home team top 6	-2.660	-1.056	-12.95**
	(4.844)	(5.511)	(6.199)
Home lose x Home team top 6	8.318	9.483	32.64***
	(5.226)	(5.855)	(8.840)
Observations	2,816	2,802	1,403
R-squared	0.417	0.395	0.384

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 *Closed* dummy.

4 Discussion

Our results confirm that the home-team bias of the referees does exist. 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 by approximately 13 seconds when the extension is in favor of them compared to the case when the extension does a disservice to them.

Moreover, we even in absence of home fans, thus it cannot be attributed to the social pressure coming from the crowd. Moreover, this bias is stronger when it supports a big, influential team, and is especially large when an influential team is losing at home to minnow. Referees thus show evidence of favoritism to hosting teams, in particular when they are influential and their opponent is not. We found stability over leagues and time, and learnt that during closed games, influential teams relatively benefited even more.

Throughout this exercise, we have discussed this gap as evidence of a personal bias driven by indirect social pressure. An alternative explanation could be that this effect may be driven by an expected quid pro quo: either direct bribery or career concerns. We believe at this level, such bribery is highly unlikely. Could referees be helping teams, especially influential ones, in the hope that they may help them professionally in return? Although still not very likely, it is a possible scenario, and one that we can partially investigate. In particular, we can look at future benefits in the form of appointments to referee at the prestigious and lucrative UEFA Champions League and UEFA Europa Leagues. We looked at the correlation between the probability of a referee getting a job at either one as a function of the their average home-team bias as well as the average influential team bias for preceding years. A suggestive evidence would show that conditioning on experience, referees with a larger big team bias would get more opportunities for European refereeing jobs. We found no evidence at all, suggesting that no visible career concerns could drive this bias.[21]

After ruling out the direct self-interested motivations from the list, we are left with some more nuanced explanations. The bias may be the consequence of an unconscious bias to conformity. People are often uncertain about decisions and in such case more willing to err on the side of conformity – which, in this case, implies helping big teams with more power. As different aspects of power (such as financial power, size of supporting base, clout outside the league, etc.) in football are often intertwined, we cannot empirically dissect whether the effect comes, in general, through money or popularity.

Moreover, risk aversion is also a possible explanation: referees make errors, including not compensating enough for wasted time. Such an error may get unpunished if it hurts a small team but may get massive press coverage in media when a large team is hurt.

[21] For details, see Section 5.5 in Appendix.

5 Appendix

5.1 Stoppage time

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 International Football Association Board, the extra 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 time, the fourth official indicates the minimum additional time. The actual extra time can be increased but not reduced by the referee.

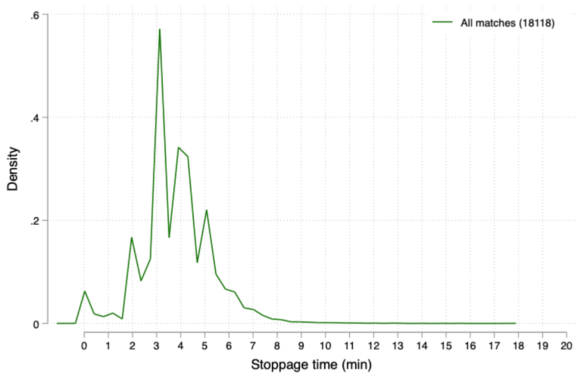
The referee decision is driven by a set of detailed rules regarding which events may generate an extension. The exact length is, however, determined by the referee.

Figure 4 plots the kernel density estimate of stoppage time distribution for these two outcomes separately (Panel 4b), along with that for the full sample (Panel 4a). 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.

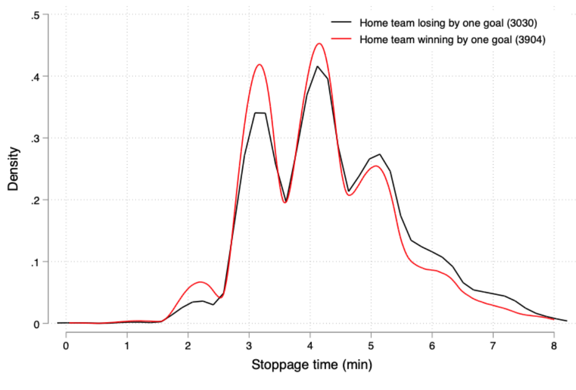
5.2 Covid and closures

Due to the Covid-19 outbreak in Europe in Spring 2020, practically every football league was suspended as of the second weekend of March. The last round before suspension was played behind closed doors in Italy on the 8th and 9th of March, as well as the last game in our sample, played on the 11th of March in Germany. The remaining games of the 19/20 season were played, with a more intensive schedule, starting 16 May in Germany, starting 11 June in Spain, starting 17 June in England, and starting 21 June in England. Each of these games were played behind closed doors, with no supporters allowed to enter the stadium. In France, remaining games of the 19/20 campaign were not played.

Figure 4: Distribution of additional time



(a) All matches



(b) Matches with one goal difference

Season 20/21 started in August 2020 in France, in September 2020 in the rest of the leagues. The vast majority of the games in the season were played behind closed doors. Depending on the severity of the Covid situation, some leagues allowed a restricted number of attendees 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 3 shows, this partial opening allowed to fill the stadiums 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%.

5.3 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.

5.4 Alternative model estimations

Table 7 presents the models of Table 2 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 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 3.

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

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: 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	10.91
	(1.94)	(1.97)	(3.49)	(3.80)	(7.38)
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***	-0.70***
	(0.32)	(0.32)	(0.14)	(0.16)	(0.16)
Away value million	-0.46***	-0.44***	-0.48***	-0.62***	-0.63***
	(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 difference	0.21	0.17	0.18	0.33*	0.32*
	(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***	225.54**
				(89.50)	(98.69)
Home lose × Attendance 1000				172.04	222.77
				(125.37)	(135.43)
Attendance/Capacity (%)					0.09
					(0.08)
Home lose × Attendance/Capacity (%)					-0.08
					(0.10)
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 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 extra 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 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 extra 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 team 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 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 extra 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 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 extra 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	10.56***	5.56***	8.05***
	(0.82)	(0.71)	(0.78)
Season FE	Yes	Yes	Yes
League FE	Yes	Yes	Yes
Home team FE	Yes	Yes	Yes
R^2	0.75	0.74	0.75
Favorite definition	bigteams	oddsdiff	valuediff
Observations	16687	6457	6457

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.

5.5 Referees and UEFA league

In the section we briefly present a simple analysis of referee's European career. The main question we address here is if more biased referees are more likely to get work at UEFA Champions League and UEFA Europa League games. There are 233 referees in our data. In our sample, on average, a referee would work in 5 seasons, on 70 games, 33 of which would have a one goal difference after 90 minutes.

We define biased in two ways: having higher average home-team bias, and having higher average influential team home-team bias. To calculate them, we aggregate first calculate the expected extension as in equation 3, and aggregate at the referee and season level. The result is a new dataset that in unbalanced panel of referees ($r = 1 \dots 233$) and seasons ($s = 2011 - 12 \dots 2020 - 21$). The total number of observation is 1148, covering on 1-10 seasons for 233 referees (on average 4.92 seasons per referee). This dataset was merged with information on referee work during this period in Champions League and the Europa League.

We first estimated a linear probability regression, the dependent variable, $euro_{r,s}$ is one when referee r worked at least a single game at either competitions in season s. It is a cross sectional regression for the 3rd season of the referee in the dataset.

$Bias_{r,s}$ 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. Experience is measured is number of gamed refereed in national leagues. Both bias and experienced is measured as the sum of values in two past seasons $((s-1, s-2)$, and α_s denotes season dummies (capturing if the referee is new). Results are presented in Table 13.

$$Pr(euro_{r,s} = 1) = alpha_s + \beta_1 bias_{r,(s-1,s-2)} + \beta_2 experience_{r,(s-1,s-2)} + \epsilon_{r,s}, \tag{6}$$

In column (1) of Table 13, We find that referees who work in more games are slightly more likely to work in European games, but average bias does not matter at all. Results are robust to alternative specifications.

Second, instead of a cross section, we looked at within referees cases: does a referee get a better chance when he is more biased than average? We now add Θ_r referee fixed effects to same model, and estimate it for all seasons.

$$Pr(euro_{r,s} = 1) = alpha_s + \beta_1 bias_{r,(s-1,s-2)} + \beta_2 experience_{r,(s-1,s-2)} + \Theta_r + \epsilon_{r,s}, \tag{7}$$

Once again, we find no correlation as shown in Column (2) of Table 13. Both these results are robust to the count of games (column (3)), or the count of games in CL only (Column (4).

Finally, we repeated both regressions, with a different bias definition. Let us consider only games when the home team is influential: $bias_influence_{r,s}$ is now the difference between the average seasonal (s) deviation but only for the perspective of influential teams. There is much fewer observations here, so we only estimated a pooled OLS with season dummies as well as variable counting the number of season a referee has worked domestically.

As shown in Table 13 column (5), once again, we find no correlation.

Table 13: Referee bias and European jobs

	(1) CL+EL	(2) CL	(3) CL+EL	(4) CL+EL	(5) CL+EL
	binary	binary	count	binary	binary
Bias	0.000288	0.000333	0.00144	0.000266	2.86e-05
	(0.000549)	(0.000371)	(0.00195)	(0.000323)	(0.000689)
Matches in league	0.0119**	0.00457	0.0471	0.00404	0.00593
	(0.00518)	(0.00528)	(0.0290)	(0.00460)	(0.00644)
Constant	-0.0861	0.0131	-0.530	0.198	0.312
	(0.195)	(0.194)	(1.030)	(0.196)	(0.308)
Observations	104	104	104	558	71
R-squared	0.163	0.079	0.149	0.740	0.256

Note: Linear probability model regressions (column 1,2,3,5), with referee fixed effects in column 4. Robust standard errors (col 1,2,3,5), referee level clustered standard errors (col 4). Predicted bias is based on 3. In columns 1-4, the bias is difference in average residual when the home team is losing vs winning. In column 5, it is difference between influential and non-influential teams when losing at home.

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

5.6 Alternative influential 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.

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.]

References

Akerlof, G. A. (1980). A Theory of Social Custom, of which Unemployment may be One Consequence. *The Quarterly Journal of Economics*, *94*(4), 749–775. Retrieved from https://doi.org/10.2307/1885667 doi: 10.2307/1885667

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In *Groups, leadership and men; research in human relations.* (pp. 177–190). Oxford, England: Carnegie Press.

Bénabou, R., and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, *96*(5), 1652–1678. doi: 10.1257/aer.96.5.1652

Bernheim, B. D. (1994). A Theory of Conformity. *Journal of Political Economy*, *102*(5), 841–877.

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

Buraimo, B., Forrest, D., and Simmons, R. (2010). The 12th man?: refereeing bias in english and german soccer. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *173*(2), 431–449. Retrieved from http://www.jstor.org/stable/40666253

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

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

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

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

DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, *127*(1), 1–56.

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

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

Gerber, A. S., Green, D. P., and Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American political Science review*, *102*(1), 33–48.

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

Klein, J. G., Smith, N. C., and John, A. (2004). Why we boycott: Consumer motivations for boycott participation. *Journal of Marketing*, *68*(3), 92-109.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, *14*(3), 115–136. doi: 10.1257/jep.14.3.115

Milgram, S. (1963). Behavioral study of obedience. *The Journal of abnormal and social psychology*, *67*(4), 371.

Nevill, A. M., Balmer, N. J., and Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of sport and exercise*, *3*(4), 261–272.

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

Rickman, N., and Witt, R. (2008). Favouritism and financial incentives: a natural experiment. *Economica*, *75*(298), 296–309.

Scoppa, V. (2008). Are subjective evaluations biased by social factors or connections? an econometric analysis of soccer referee decisions. *Empirical Economics*, *35*(1), 123–140.

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.