The Not-So-Beautiful Game

Alex Peterson & Ben Muehleisen

2023 Senior Thesis

Department of Economics

Pomona College

I.   Introduction

"Football is nothing without fans." These simple words spoken by Jock Stein underpin the importance that fans play in a football match, and that truth was never more evident than in the "ghost" matches played during the recent years due to COVID-19. The silent and empty stadiums felt hollow without the tens of thousands of fans singing, chanting, and cheering. The loss of that support provided an opportunity to analyze the effect that the fans in stadiums had on the game. Most people are familiar with the concept of "home-field advantage" where the home team tends to win more because of the advantage their fans give them. But how exactly are the home fans influencing the game? There is the psychological element of having thousands of fans supporting one team and trying their best to disrupt the other which may motivate players to play better because of the energy it can give them. On the other hand, it may place undue pressure on the players, making them play more conservatively because of the fear of making a mistake, which results in deafening boos and piercing whistles. Of the many impacts that the presence of fans has on the game of football, one of the most controversial aspects is their influence on referees. As a part of a game that is meant to be unbiased, referees are constantly being judged and criticized for the decisions they make, especially when they rule in favor of the home team, because of the perception that a supposedly neutral third party is being affected by a biased home crowd.

If a contentious tackle is made by the away team, even if the referee initially thought it was within the laws of the game, they might have second thoughts if an entire stadium of angry fans disagrees. Especially in a low-scoring sport like football where a small decision can have a monumental impact on the game, fans are particularly keen to make their thoughts and complaints heard at every opportunity.

This paper seeks to explore the relationship between fans in the stadium and disciplinary decisions being made by referees in football matches before, during, and after the global COVID-19 pandemic. Due to the unique circumstances of the times, with fans not being allowed in stadiums, we will take advantage of this natural experiment and look at the rate of fouls and yellow cards being given to both home and away teams. Comparing different seasons with and without fans will allow us to see if there is a significant effect that fans are making on players' and referees' decisions.

II.   Literature Review

There is an extensive existing literature on home advantage, referee bias, and the effectiveness of tactical fouling in professional football, however, there is very little research exploring the impact that fans have on referee decision-making. This is likely due to the fact that in the age of modern stat keeping, fans have always been present in the stadiums of the top-flight European football leagues. As each of the teams play each other twice - once on each team's home pitch - researchers could compare home and away variables to make observations. However, now the COVID-19 pandemic has given us a unique opportunity to observe and use data in which no fans - or a limited number of fans - were present in stadiums.

The work of Michael Christian Lietner and Fabio Richlan, in their 2021 paper, "No Fans – No Pressure: Referees in Professional Football During the COVID-19 Pandemic," is the closest existing study to what we are looking to uncover through our research. Lietner and Richlan took data from matches played in the top leagues of Spain, England, Germany, Italy, Russia, Turkey, Austria, and the Czech Republic, during both the 2018-2019 regular attendance season and the 2019-2020 no fan attendance, "Ghost Games," and focused on the game result, fouls, bookings,

and reasons for bookings. They used two-sided Mann-Whitney-U-tests – a method of measuring central tendencies between two groups – for independent data, along with four different tests for dependent data that were used to investigate a potential advantage effect. They came to two main findings: 1) Home teams were booked significantly more often with yellow cards for committing fouls in ghost games than in regular games (Lietner and Richlan, **Figure 4**), and 2) The home advantage effect in every respective league vanished in the ghost games of the 2019-2020 season. In the 2018-2019 season, teams had "a quantified home performance of 63.5% home-win-rate," and during the ghost games "home performance diminishes to a 52.5% home-win-rate" (6). They find that referees don't directly disadvantage the away team, because there was no significant change in yellow cards given to away teams between the regular attendance season and the ghost games, but rather that referees unconsciously give an advantage to home teams in regular attendance games. They conclude that referees are able to act more objectively when determining whether or not to take disciplinary action through a booking without the presence of fans.

Differences (Total numbers) in Yellow Cards between Season 2018/19 & 2019/20 (ESP, ENG, GER, ITA, RUS, TUR, AUT, CZE)

Reason for Diciplinary Action	2018-2019 Season		2019-2020 Season		Ghost Games in 2019-2020 Season	
	Home Teams	Away Teams	Home Teams	Away Teams	Home Teams	Away Teams
Critisicm	77	93	118	127	41	34
Unfair Sportsmanship	229	249	97	114	-132	-135
Foul	910	1013	1148	1041	238	28

FIGURE 4 | Absolute differences in Criticism, Unfair Sportsmanship, and Foul Play. Table shows yellow cards (crisicism, unfair sportsmanship, and foul play) awarded to home and away teams in the top leagues of Spain, England, Germany, Italy, Russia, Turkey, Austria, and Czech Republic. Based on aggregated match results statistics of rounds played in the leagues as ghost games (no or strongly limited attendance) in season 2019/20 and respective round of regular matched (regular attendance) in season 2018/19 N = 146).

It is worth noting from Lietner and Richlan's Figure 4 that yellow cards for unfair sportsmanship dropped dramatically between the regular attendance season and the ghost games, which indicates that the presence of fans may also influence the players' aggression.

Our work will build off of Lietner and Richlan's research, and look to uncover more findings while checking to see if we draw similar conclusions. Our study will benefit from the passing of time, as we now have data for the entire 2020-2021 season which was also impacted by COVID-19, as well as data for the 2021-2022 season where fan attendance returned to normal. We are hopeful that having far more observations and focusing only on the top 5 European leagues will offer us the most accurate insight into the influence fans have over referees.

III.   Data

This paper will use data from Opta, a company that records detailed event data from football games across the biggest leagues going back over 20 years to the 1999/2000 season. It contains all the variables necessary for our analysis, including season possession, fouls, and yellow cards, as well as distinguishing between home and away matches. Along with fouls and yellow cards, red cards are the only other decision the referee will make in a game. However, we do not look at red cards as they are too rare of an occurrence in matches to get consistent data due to their randomness and infrequent occurrences. Furthermore, red cards are a VAR-reviewable event, which means that although the on-field referee will still have the final decision, VAR can intervene to correct any mistakes, which would hypothetically reduce the impact of home-crowd advantage as any video review is done outside stadiums and the influence

of fans, the very effect we are trying to look at. It would additionally be difficult to account for the effect that VAR influence may have on any analysis of referee bias that we conduct.

Our dependent variables are normalized yellow cards and fouls per game to measure the impact of fans on the referee's decisions. Independent variables include the time period, the league, and whether the matches were played at home or away.

We need to account for the time of possession the team had or did not have since both yellow cards and fouls are defensive metrics. Naturally, a team that possesses the ball 60% of the time will on average commit fewer fouls and yellow cards than a team that possesses the ball 30% of the time. This is because the first team has to defend for a significantly smaller portion of the game, leading to lower averages. To account for this, we use two new variables: Normalized Yellow Cards per Game (Norm_YCPG), and Normalized Fouls per Game (Norm_FPG). Norm_YCPG is finding the number of yellow cards the team would have committed if they had 90 full minutes of nonpossession. Essentially it is extrapolating the rate at which the team is receiving yellow cards per their level of possession and adjusting it to no possession of the ball whatsoever. The same applies to Norm_FPG for fouls instead of yellow cards. Their formulas are as follows:

a) $Norm\_YCPG = \dfrac{Yellow\ Cards\ per\ Game\ (YCPG)}{1 - \frac{Percent\ Possession}{100}}$

b) $Norm\_FPG = \dfrac{Fouls\ per\ Game\ (FPG)}{1 - \frac{Percent\ Possession}{100}}$

With these two variables, we can compare each team's respective numbers for fouls and yellow cards without worrying about different teams' playing style or philosophy, and how much time they spend with or without the ball. One consideration we must take is that this increases the total number of fouls/yellow cards across a season as the denominator will always be less than one. Therefore, the "Normalized" total will be greater than the original tally. In the case of

"Norm_YCPG," it is finding the number of yellow cards per game if that team was out of possession (on the defensive) the entire match, so the number would be higher than the original statistic. These new figures allow for greater comparison across each time period, the relationship that we are attempting to analyze, which is why we chose to normalize the statistics based on possession. Furthermore, after normalization, taking the "per game" numbers instead of season totals ensures that the short affected portion of the 2019-2020 season can be included, as well as accounting for the different numbers of total games played across the different leagues. A Bundesliga team will only play 34 games in a season as opposed to the other four leagues, all of which have 38 games in a season.

We will additionally look at and control for five different leagues for this study, representing the most popular and biggest football leagues in the world. They were chosen due to the highest number of average fans in stadiums, so the effect of their absence would be more pronounced. These are the Bundesliga (Germany), La Liga (Spain), Ligue 1 (France), Premier League (England), and Serie A (Italy). Furthermore, inconsistent COVID measures regarding fan presence in stadiums along with the decreasing standard of quality of referees as you move further away from the top leagues would make it harder to draw inferences from the results. This is why we restricted our analysis to the top five football leagues. We control for any effect the league may have on the data by creating four indicator variables representing each of the leagues, with the Premier League as our reference case.

This paper will also analyze periods in which fans are present and not present, to draw the most comprehensive conclusions and build off of Lietner and Richlan's work. Our data consist of three distinct time periods: "Pre-COVID" consists of data starting from the 2014-2015 season through the end of the 2018-2019 season, "COVID" consists of the portion of the

2019-2020 season that was disrupted by COVID-19, as well as the entirety of the affected

2020-2021 season, and "Post-COVID" consists of the 2021-2022 season. The Pre-COVID time

period serves as our foundational understanding of data in a time of "normalcy" in which fans

are present, and stadiums are at full capacity in the top five leagues. The COVID time period

contains our data during which there were either no fans present (end of 2019-2020 season), or

only a very limited capacity of fans (at times in the 2020-2021 season). We split this period into

two periods ("COVID-a" and "COVID-b"), in which COVID-a isolates the end of the 2020

season data, and COVID-b contains both the end of the 2020 season data as well as the entire

2020-2021 season data. This was in an effort to examine the COVID-19-impacted data as a

whole and just the last few match days of the 2020 season when the COVID-19 disruption was at

an extreme. While the 2020-2021 season was certainly affected by the ongoing pandemic

limiting fan attendance, towards the end of 2021, different leagues opened up the stadiums to

host a very low capacity of fans for match days. Post-COVID consists of data during which fans

have fully returned to the stadiums, bringing us back to a time of "normalcy". Our goal is to

draw conclusions on the impact that fans have on referee decisions by focusing on the

Pre-COVID, COVID-a, and COVID-b time periods, but we will also look at the Post-COVID

time period to determine if there are any noticeable lasting impacts of the COVID-19 disruption.

We created dummy variables representing each time period, using Pre-COVID as our reference

level.

  Matches being played home/away is the last variable we accounted for. This is the main

variable of interest along with its interactions to show how a home crowd is more or less likely to

influence referee decisions in the game. **Table 1** shows a summary of all the variables we

analyzed, including the raw figures for yellow cards and fouls, along with the average season

possession of each team. Average season possession is equal to the total percentage of possession

for each team divided by the number of games they play over the course of the season.

Table 1: Summary Statistics

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Yellow Cards	34.46	35.00	12.94	0	92.00
Norm_YC	69.19	69.10	25.81	0	173.16
Norm_YCPG	3.78	3.73	1.23	0	9.11
Fouls	224.10	235.00	64.42	27.00	355.00
Norm_Fouls	450.62	474.55	129.54	64.67	720.93
Norm_FPG	24.70	25.60	5.84	6.02	39.10
COVID-a_Dummy (Empty Stadiums)	0.10	0	0.30	0	1
COVID-b_Dummy (Partially Full Stadiums)	0.13	0	0.33	0	1
Post-COVID_Dummy (Full Stadiums)	0.13	0	0.33	0	1
Away/Pre-COVID	0.32	0	0.47	0	1
Away/COVID-a	0.05	0	0.22	0	1
Away/COVID-b	0.06	0	0.25	0	1
Away/Post-COVID	0.06	0	0.25	0	1

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Ligue1_Dummy	0.18	0	0.39	0	1
SerieA_Dummy	0.21	0	0.41	0	1
LaLiga_Dummy	0.21	0	0.41	0	1
Bundesliga_Dummy	0.19	0	0.39	0	1
Possession (%)	49.96	49.2	5.40	20.6	68.2

Note: Observations are 1528 for all variables. All yellow cards and fouls data describes the team that concedes these disciplinary actions.

IV. Methodology

The research objective of this paper is to examine the relationship between the influence of fans in a stadium on the decisions a referee makes during the game. Whether the official will give an unconscious bias towards the home team, giving them fewer yellow cards and fouls, because of the home fans yelling and screaming at them. In order to do so, we constructed two multivariate regression models, both of which were estimated using ordinary least squares (OLS). The full models are shown in the appendix; we account for each of the variables and interactions as explained later in the section.

The interactions between home/away matches and the time period variables were the only interaction effects we included in this study. Initially, we included all two-level interactions between the variables: home, league, and time period. However, we performed an F-test for significance at a 5% level of significance to assess the contribution of the interaction terms between league and home/away matches in the model. The F-test revealed that the p-value for

the interaction terms was above the significance level, indicating that the interaction terms did not significantly contribute to the model. Based on these results, we decided to drop the interaction terms between home/away and league from the final model, as the smaller nested model provided a better measure of the results. Additionally, we wanted to exclusively look at the effect of each time period on referee decisions, so the interaction of the effect of home support within each time period is relevant and important to the question we are trying to answer. We only want to control how much each league affects the referee's decisions regarding yellow cards and fouls.

In total, we conducted two different regressions. Within each time period, we were looking to see if away teams had significantly higher rates of yellow cards/fouls per game, as well as whether there was an association between each league and the number of yellow cards/fouls given out across both home and away games. We can then compare each of the time periods and the associated correlation between the variables we found within the models to see how fans in the stadium affected referee decisions with regard to yellow cards and fouls per game.

V.   Results

Table 2. Regression Summary	Normalized Yellow Cards Per Game Observations: 1528 / R^2: 0.57		Normalized Fouls Per Game Obervations: 1528 / R^2: 0.67	
	Coefficients	P-value	Coefficients	P-value
Intercept	3.03	<0.001	21.86	<0.001
COVID-a_Dummy (Empty Stadiums)	-1.90	<0.001	-13.53	<0.001
COVID-b_Dummy (Partially Full Stadiums)	-0.03	0.740	-0.49	0.188
Post-COVID_Dummy (Full Stadiums)	0.19	<0.001	-1.15	0.002
Away/Pre-COVID	0.59	<0.001	0.94	<0.001
Away/COVID-a	-0.15	0.244	-0.38	0.481
Away/COVID-b	0.22	0.058	-0.31	0.514
Away/Post-COVID	0.44	<0.001	0.36	0.446
Ligue1_Dummy	0.31	<0.001	4.02	<0.001
SerieA_Dummy	1.21	<0.001	6.42	<0.001
LaLiga_Dummy	1.57	<0.001	5.32	<0.001
Bundesliga_Dummy	0.31	<0.001	5.03	<0.001

Table 2 outlines a summary of our regressions on our outcome variables of Normalized

Yellow Cards per game and Normalized Fouls per game. Normalizing these variables allowed us

to measure the amount of yellow cards received and fouls committed per 90 minutes of

non-possession. This was important in order to have accurate measurements, because with the

exception of yellow cards given for unsportsmanlike conduct or dissent, the vast majority of

yellow cards received and all fouls committed occur on defense. So top teams who maintain the

majority of possession across the 90 minute game will likely receive fewer yellow cards and

commit fewer fouls than their opponents. Our regressions' R-Squared values for Normalized

Yellow Cards Per Game and Normalized Fouls Per Game are fairly strong at 0.57 and 0.67

respectively.

Our results offer some interesting directly relevant takeaways for the purpose of our

paper, as well as some general observations. The Away/Pre-COVID, Away/COVID-a, and

Away/Post-COVID variables offer the most directly relevant takeaways, as they give us insight on the changes of the home advantage effect across the time periods. We see that across all leagues, away teams in the Pre-COVID time period receive 0.59 more yellow cards and commit 0.94 more fouls per 90 minutes of non-possession than home teams, and this was statistically significant. This corroborates previous literature that has found the home advantage effect to be real. However, during the COVID-a time period we find no statistical significance of away teams being at any disadvantage to home teams in terms of yellow cards and fouls. This is what we were expecting to find, and this indicates that referees called a more even game during the height of COVID's impact on Europe's top five football leagues. We also find no statistical significance of a home advantage effect during the broader COVID-b time period, however it is closer to significance at the 5% level than the COVID-a time period. This is likely due to the COVID-b time period including games in which some stadiums allowed a limited capacity of fans to attend in the 2020-2021 season. As we would expect, in the Post-COVID time period, we see the home advantage effect return in terms of away teams receiving 0.44 more yellow cards per 90 minutes of non-possession than home teams. Strangely, we did not find statistical significance here for away teams committing more fouls per 90 minutes of non-possession than home teams.

In terms of some general observations, we see that across all leagues, home teams in the COVID-a time period received 1.9 fewer yellow cards and committed 13.53 fewer fouls per 90 minutes of non-possession than home teams in the Pre-COVID time period. This could indicate that players are less aggressive without the atmosphere that their home fans create. While this was significant in the COVID-a time period with p-values < .05, we did not see significance for the COVID-b time period in this regard. Perhaps this is due to various stadiums returning to limited capacity in the 2020-2021 impacted season. As we would expect, home teams during the

Post-COVID time period essentially received the same amount of yellow cards and committed the same amount of fouls per 90 minutes of non-possession as home teams during the Pre-COVID time period. When looking at our league variables, we see that every league receives more yellow cards and commits more fouls per 90 minutes of non-possession than the Premier League, and all of these variables were statistically significant. This was surprising to see as the football community generally views the Premier League as the most aggressive and rough of the top five leagues. Our results suggest that perhaps the Serie A or La Liga should take this title, with teams receiving 1.21 and 1.57 more yellow cards per game respectively, and committing 6.42 and 5.42 more fouls per game respectively than Premier League teams. Although, our results are not a conclusive indicator of aggression as this could also be due to different refereeing styles between the leagues. If referees in the Serie A and La Liga call a tighter game than those in the Premier League we would make the same observations.

VI.   Conclusion

Our goal of this paper was to observe how fans' presence or absence in the stadiums of the top five European football leagues impact referee disciplinary decisions, specifically focusing on yellow cards and fouls. We took advantage of a natural experiment set-up due to COVID-19, where we had data from the 2014-2018 seasons to create a Pre-COVID time period, the impacted seasons of 2019-2020 and 2020-2021 to create a COVID time period (broken into COVID-a and COVID-b), and the season of 2021-2022, in which fans had completely returned to the stadiums, to create a Post-COVID time period.

We looked to build upon the earlier research of Lietner and Richlan in their 2021 paper, "No Fans–No Pressure: Referees in Professional Football During the COVID-19 Pandemic". As

we found no statistical significance of away teams receiving more yellow cards and committing more fouls per 90 minutes of non-possession in the COVID-a and COVID-b time periods, our results agree with Lietner and Richlan's primary finding that the home advantage effect vanished during the COVID-19 impacted seasons. We agree with Lietner and Richlan that the most plausible explanation for this is because referees must have a subconscious bias towards home teams caused by pressure from home fans. Our data and results provide additional insights to that of Lietner and Richlan's, primarily because of our ability to isolate the top five European leagues for consistency of level of play and referee standards, as we had the benefit of having data on an additional season that was disrupted by COVID-19. In addition, we think that our decision to normalize our outcome variables separates us from previous literature as yellow cards and fouls occur from defensive actions. So, it was important to normalize for yellow cards received and fouls committed per 90 minutes of non-possession.

Something that we did not control for that may have had a small influence on our results in some way is the top five European leagues temporarily deciding to increase the allotted number of subs from three to five during the COVID-19 impacted seasons. This has since become permanent in efforts of protecting players from injury as many teams play three of four times a week. In addition, two one-minute water breaks (one each half) were also introduced during this time. We don't think that this should have a huge impact on our results, but perhaps teams having players on the pitch with fresh legs could change the amount of yellow cards received and fouls committed. In addition, coaches can sub off players that have already received a yellow card, and replace them with players that have a clean slate and may be less fearful of conceding a disciplinary action. Overall, we are pleased that our research corroborated the main findings of Lietner and Richlan and offered some other interesting and unexpected observations.

Appendix

Model 1: Effect on Normalized Fouls per Game

$Norm\_FPG = \beta_0 + \beta_1\ COVID\text{-}a\_Dummy_i + \beta_2\ COVID\text{-}b\_Dummy_i + \beta_3$

$Post\text{-}COVID\_Dummy_i + \beta_4\ Away/Pre\text{-}COVID_i + \beta_5\ Away/COVID\text{-}a_i + \beta_6\ Away/COVID\text{-}b_i + \beta_7$

$Away/Post\text{-}COVID + \beta_8\ Ligue\ 1\_dummy_i + \beta_9\ Serie\ A\_dummy_i + \beta_{10}\ LaLiga\_dummy_i + \beta_{11}$

$Bundesliga\_dummy_i + \varepsilon_i$

Model 2: Effect on Normalized Yellow Cards per Game

$Norm\_YCPG = \beta_0 + \beta_1\ COVID\text{-}a\_Dummy_i + \beta_2\ COVID\text{-}b\_Dummy_i + \beta_3$

$Post\text{-}COVID\_Dummy_i + \beta_4\ Away/Pre\text{-}COVID_i + \beta_5\ Away/COVID\text{-}a_i + \beta_6\ Away/COVID\text{-}b_i + \beta_7$

$Away/Post\text{-}COVID + \beta_8\ Ligue\ 1\_dummy_i + \beta_9\ Serie\ A\_dummy_i + \beta_{10}\ LaLiga\_dummy_i + \beta_{11}$

$Bundesliga\_dummy_i + \varepsilon_i$

References

Fischer, K., and Haucap, J. (2020). Does Crowd Support Drive the Home Advantage in Professional Soccer? Evidence From German Ghost Games During the COVID-19 Pandemic.

Goumas, Chris. "Home advantage and referee bias in European football." European journal of sport science vol. 14 Suppl 1 (2014): S243-9. doi:10.1080/17461391.2012.686062

Leitner MC and Richlan F (2021) No Fans–No Pressure: Referees in Professional Football During the COVID-19 Pandemic. Front. Sports Act. Living 3:720488. doi: 10.3389/fspor.2021.720488

Phatak, Ashwin A et al. "The Dirty League: English Premier League Provides Higher Incentives for Fouling as Compared to Other European Soccer Leagues." Journal of human kinetics vol. 80 263-276. 31 Oct. 2021, doi:10.2478/hukin-2021-0095

Sors, Fabrizio et al. "A complete season with attendance restrictions confirms the relevant contribution of spectators to home advantage and referee bias in association football." PeerJ vol. 10 e13681. 4 Jul. 2022, doi:10.7717/peerj.13681

Sors, Fabrizio et al. "The sound of silence in association football: Home advantage and referee bias decrease in matches played without spectators." European journal of sport science vol. 21,12 (2021): 1597-1605. doi:10.1080/17461391.2020.1845814

Wunderlich, Fabian et al. "How does spectator presence affect football? Home advantage remains in European top-class football matches played without spectators during the COVID-19 pandemic." PloS one vol. 16,3 e0248590. 31 Mar. 2021, doi:10.1371/journal.pone.0248590