A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

WHEN THE 12TH MAN IS GONE:

HOW FANS AFFECT THE PERFORMANCE OF DIFFERENT PLAYER GROUPS

BRUNO MOREIRA

Work project carried out under the supervision of:

Alex Armand

15/12/2022

Abstract

Using the unique circumstances provided by the COVID-19 pandemic, I investigate the effects of fan absence in stadiums on the performance of different football player groups in England. During the 2019/20 season, most non-British Isles player groups perform worse without fans. In the following two seasons, I find evidence of a reduced home advantage in ghost games and proof that English players perform significantly better without fans, especially in away games. Regarding the latter, further investigation suggests that, specifically players that are not former academy players of their team perform better without fans than those who played for their current team at the youth level.

Keywords: Fixed Effects; COVID-19; Football; Performance.

Acknowledgements

I want to thank my supervisor, Prof. Alex Armand, for the guidance and advice he has provided throughout my thesis. I have been fortunate to have a supervisor who was so invested in my work and responded quickly to my questions and doubts. I am also grateful to Adeel Tallat and Matthias Piot for their assistance in the data acquisition and cleaning. Finally, I thank Mauro Caselli for his valuable comments.

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1 Introduction

Football is widely recognised as the most popular sport globally because it is a sport for everyone. It can entertain and unite people regardless of origin, making it one of the most inclusive sports. Another reason for the massive popularity of football, also referred to as "The Beautiful Game", is that it is visually appealing. We enjoy watching football because of strikes into the top corner, flawless slide tackles or perfectly timed through balls, but we also love the spectators' fever and clamour. In football, the collective term 12th man is used to refer to the fans in the stadium, implying that they have a potentially valuable and significant role in the game. Spectators often create loud sounds or chants to encourage their team; or distract or demoralise the opposing team. In his autobiography *My Turn: The Autobiography*, Johan Cruyff reveals what it was like playing at Anfield, home of Liverpool FC: "There's not one club in Europe with an anthem like "You'll Never Walk Alone." There's not one club in the world so united with the fans. I sat there watching the Liverpool fans and they sent shivers down my spine. A mass of 40,000 people became one force behind their team" (Cruyff 2016, 36).[1]

The question is how the presence of fans impacts the performance of players. On the one hand, football players may thrive when feeding off the energy and electricity provided by a crowd. On the other hand, the lack of fans may relieve the anxiety or pressure that football players may experience, which could improve their performance. Finding an answer to this question would help to clarify how the social environment affects the efficiency of economic agents.

The COVID-19 pandemic provides the perfect possibility to gain an understanding of the effects of fan presence on the performance of football players. The Premier League (the highest tier of the men's English football pyramid) suspended its 2019/20 season on March 13

[1] Johan Cruyff (1947 – 2016) was a Dutch professional football player and manager. He is recognised as one of the greatest players in football history, as well as one of the best managers ever.

(2020) in the aftermath of the outbreak of the COVID-19 pandemic. When the Premier League resumed later in June, the remaining games were played behind closed doors i.e., without a crowd watching. By the means of this natural experiment, I can determine whether Premier League players' performances significantly change when the 12th man is gone.

My key result for season 2019/20 is that almost all non-British Isles player groups perform worse in the post-lockdown games than in the pre-lockdown games.[2] A possible explanation that is rather difficult to test is that these country groups may have struggled more with the challenges that came with the first lockdown, e.g. travel restrictions which affected non-UK players more since the possibilities of seeing their families were reduced. Further investigation shows that most non-English subgroups perform worse in away games without fans than before the first lockdown. Non-English country groups also perform significantly worse in away games without fans when compared to English players.

My second key result is that in the following two seasons (season 2020/21 was mostly played without fans, and season 2021/22 was entirely played with fans), on average, players perform worse in home games without fans than in home games with fans, while the opposite result holds in away games. In particular, English players experience a significant increase in their performance in away games without fans. Further tests reveal that specifically English players that are not former academy players of their team perform better in away games without fans than English players who played for their current team at the youth level.

These effects are documented using data on players' performance derived from a distinctive, comprehensive statistical algorithm, calculated in real-time throughout the game. The analysis accounts for various time-varying elements, such as the weather during each game and individual, team, and match fixed effects.

[2] I use the term "British Isles" to refer to Ireland, Northern Ireland, Scotland, England and Wales.

Academic research on the implications of the COVID-19 pandemic on football has been quite extensive. However, since the main focus of this study lies on the players' performance, the first question should be how to measure performance. A growing number of studies have already documented why one should favour ratings based on an algorithm instead of newspaper ratings. Principe & van Ours (2021) show that newspaper ratings of professional football players suffer from racial bias. They find that conditional on their objectively measurable performance black players are awarded lower subjective ratings than non-black players. Not only can these subjective player ratings suffer from racial bias but also outcome bias. Kausel et al. (2018) argue that a penalty shootout win is connected to higher subjective performance ratings. This result persists even after one removes the players that participated in the shootout. Previous studies have documented the impact of the absence of fans on the home advantage in football. Destefanis et al. (2022) found that in the 2020/21 season, efficiency (offensive as well as defensive) increased considerably for away games, providing evidence of a generalised reduction in the home advantage. Leitner & Richlan (2021) also found that the home teams' loss percentage increased significantly, whereas the away teams' win percentage increased. In the same vein, McCarrick et al. (2021) documented a remarkable decrease in goals scored and points gained during the pandemic, which demonstrated the home team's inferior performance.

Caselli, Falco & Mattera (2020) used this opportunity that arose with the COVID-19 pandemic to test how Serie A players respond to the absence of fans. They find that African players, who are often victims of racial discrimination, receive higher objective ratings when fans are no longer present. My work investigates explicitly how Premier League players react to the absence of fans. In addition, instead of only focusing on the 2019/20 season, I also take into account the following two seasons (2020/21 and 2021/22), which allows me to test how players reacted to the absence of fans as well as to their comeback to the stadiums. Furthermore, I differentiate between home and away games.

2 Context of the study

Before the COVID-19 pandemic, football and packed stadiums went hand-in-hand. During the 2018/19 season, the last season before the COVID-19 pandemic to be entirely played with fans, the average attendance across all Premier League games was 38,168. With such a great number of fans in stadiums, it is undeniable that they can play a major role in their team's performance. It is commonly accepted that when fans cheer, it has a positive impact on their team's performance, which is why the concept of "home advantage" exists: it explains the benefit that the home team is considered to have over the away team, primarily as a result of the psychological effects that fans have on the players and referee. However, certainly not all players react the same way to the presence of fans, and the positive role fans are meant to play can sometimes be questionable.

During the 2019/20 season, numerous players encountered racist abuse by fans within the stadium. One game that stood out was between Tottenham Hotspur and Chelsea in December 2019 which had to be paused because Tottenham fans were making monkey noises at Chelsea defender Antonio Rüdiger. During the same game, a Chelsea fan was arrested following an allegation of racial abuse towards Tottenham forward Son Heung-min. Premier League fans also have a history of booing their own players. In October 2019, Arsenal's Granit Xhaka, subjected to continued abuse from the club's fans during that season, was heavily booed after being substituted against Crystal Palace. Before ripping off his shirt and vanishing into the tunnel, Xhaka replied by waiving and sarcastically cupping his ear to the supporters.

The COVID-19 pandemic provides a rare opportunity to examine how fan presence affects the performance of different player groups. On 13 March 2020, the Premier League suspended its 2019/20 season following the outbreak of the COVID-19 pandemic. The 2019/20 season resumed on 17 June, but the remaining games took place behind closed doors. The 2020/21 season started the same way the 2019/20 season finished, with games being played

behind closed doors. This applied essentially for most of the season, except for a brief spell in December 2020, during which a limited number of fans were allowed in certain stadiums, and at all grounds towards the end of the season in May 2021. The 2021/22 season was played without any restrictions on the number of fans in stadiums. Using these three seasons, I can study players' reactions to the fans' absence, as well as to their return to the stadiums. Figure 1 visually represents the COVID-19 events and their implications on the Premier League.

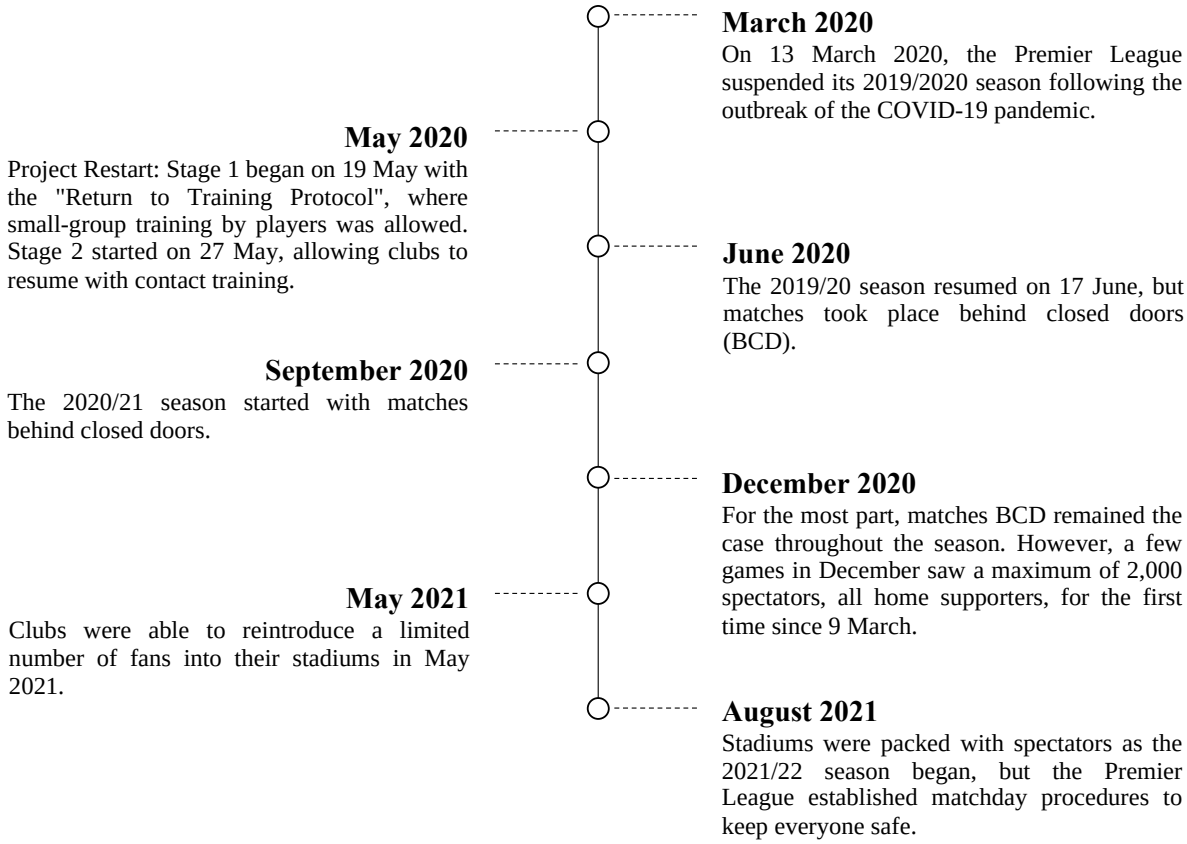
3 Data and empirical strategy

3.1 Data

The dataset was constructed from scratch using various information sources. The first step was to find a metric that objectively measures players' performance. For this purpose, I used the WhoScored.com website, which assigns ratings to each football player for every game played throughout the season.[3] Ratings scale from 0 to 10, with the highest score being 10. At the beginning of each game, each player starts with a rating of 6, which is the average. In the world of football, these ratings are regarded as the most precise, reputable, and well-known performance indicators. They are founded on a unique, extensive statistical algorithm that is calculated in real-time throughout the game. A player's rating is determined by taking into account more than 200 raw statistics and weighting them based on their impact on the game. As a result, these ratings are free of any subjectivity, often found in match reviews published in newspapers, allowing me to eliminate any source of bias. Figure 2 plots the Premier League 2019/20 average ratings per game.

[3] The data had to be scraped from the WhoScored.com website and cannot be downloaded directly.

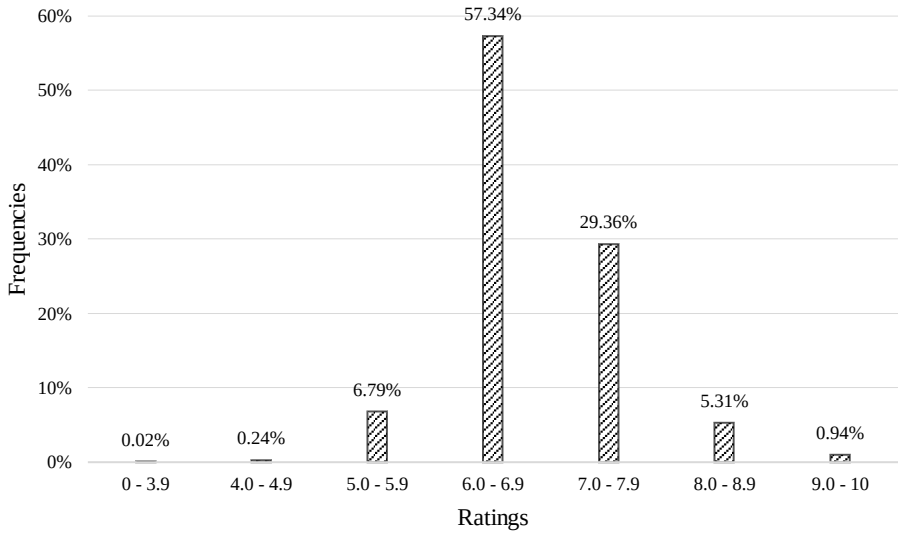
Figure 1: COVID-19 timeline



The analysis is split into two parts to create two opposite cases. The first one only focuses on the 2019/20 season as I want to analyse how players' performances changed when fans were no longer allowed in stadiums. In this part, I only include players that have at least played one game with the same team before the Premier League suspended its 2019/20 season on March 13 and after the season resumed on June 17. The second one focuses on the 2020/21 and 2021/22 seasons because, in this case, I want to analyse how players' performances changed when fans were allowed back in stadiums. I merge these two seasons to create the opposite case of season 2019/20, and similarly to the first part, I only include players that have played at least one game with the same team with and without fans in stadiums.[4] The final sample only consists of players that appeared in both cases, including 276 players across 21 teams.

[4] Season 2020/21 was mostly played without fans while season 2021/22 was played entirely with fans.

Figure 2: Premier League 2019/20 average ratings per game



I used the FootyStats.org website to obtain player-specific information such as nationality, age etc.[5] The next step was to allocate each player to one of the seven groups: England, British Isles (excluding England), European Union, South-eastern and Eastern Europe (excluding the EU), Latin America, Africa, Other. English players make up the biggest group with 38%, followed by EU players with 30.4%.

Furthermore, I complement the dataset with various weather variables using the exact time and coordinates of the stadium for each game.[6] These variables include temperature, relative humidity, windspeed, windchill, and fog, as well as a general overview of the weather state.

Table 1A focuses on the first period and displays mean performances by country group and match type (fans and no fans), as well as the corresponding t-tests to determine if there is a significant difference between the two means. Table 1B displays the same information for the second period. In Table 1A, three country groups have differences in means that are statistically significant at a 5% level. African players experienced the most significant drop in performance

[5] The data is only accessible through a premium subscription.
[6] The data was purchased at www.meteomatics.com, a specialist weather data provider, and is therefore not publicly available.

when the Premier League resumed after the lockdown. In Table 1B, only one country group has a difference in means that is statistically significant: players from the British Isles, excluding England, performed worse in the period 2 games played without fans as compared to the games played in the presence of fans. Another interesting result is that English players and South-eastern and Eastern players have negative differences in means, but none of the two is statistically significant. A common result in both cases is that players, in general, performed better with fans than they did without fans, even though the difference is minimal.

3.2 Empirical strategy

My empirical strategy to estimate the impact of fans on the performance of different player groups can be summarised by the following equation:

(1) $ln(rating)_{itm} = \alpha + \beta nofans_m + \sum_{j=1}^{J} \gamma_j nofans_m \times group_{ij} + \delta \mathrm{X}_m + \eta_{it} + \epsilon_{itm}$

where $ln(rating)$ is the natural logarithm of the rating of player i in team t for match m;[7] $nofans_m$ is a dummy variable equal to 1 if there were no fans present in the stadium; $group_{ij}$ is a dummy equal to 1 if player i belongs to player group j (f. ex. England, Africa, Latin America, etc.); X_m includes match variables that are specific to each game (f. ex. weather conditions, whether the game was a local derby or not, a home dummy, etc.); η_{it} represents a player fixed effect which controls for observable elements (f. ex. age, salary, senior experience) and unobservable features such as ability. ϵ_{itm} represents the idiosyncratic error term.

[7] A skewness test for normality allowed me to conclude that the variable $rating$ is not normally distributed. Using logs transforms the skewed variable to approximately conform to normality. Furthermore, the results do not significantly change when the analysis is conducted in levels.

Table 1A: Players' performance by country group in period 1

	Fans (1)		No Fans (2)		Difference (1) – (2)
Country group	N	Mean (SE)	N	Mean (SE)	t-test
England	1,985	6.782	758	6.775	0.007
		(0.016)		(0.027)	
British Isles (ex. England)	558	6.787	222	6.672	0.115**
		(0.026)		(0.048)	
EU	1,647	6.821	607	6.781	0.040
		(0.018)		(0.030)	
South-eastern and Eastern Europe	107	6.806	47	6.616	0.189*
		(0.061)		(0.084)	
Latin America	548	6.880	182	6.738	0.142**
		(0.033)		(0.058)	
Africa	378	7.011	144	6.808	0.203**
		(0.043)		(0.065)	
Other	204	6.860	88	6.719	0.141
		(0.055)		(0.073)	
Total	5,427	6.824	2,048	6.758	0.065***
		(0.010)		(0.016)	

Notes: Table 1A displays the number of observations, mean performance and standard error in brackets by country group and match type (with or without fans). "Fans" refers to all games played with fans in the stadium during period 1. "No Fans" refers to all games played without fans at the stadium during period 1. Period 1 refers to season 2019/20. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 1B: Players' performance by country group in period 2

	Fans (1)		No Fans (2)		Difference (1) – (2)
Country group	N	Mean (SE)	N	Mean (SE)	t-test
England	2,162	6.761	2,342	6.774	-0.013
		(0.016)		(0.015)	
British Isles (ex. England)	427	6.799	598	6.649	0.151***
		(0.032)		(0.026)	
EU	1,780	6.823	1,855	6.801	0.022
		(0.017)		(0.017)	
South-eastern and Eastern Europe	108	6.595	104	6.600	-0.005
		(0.053)		(0.058)	
Latin America	564	6.794	535	6.728	0.066
		(0.030)		(0.030)	
Africa	505	6.901	440	6.847	0.053
		(0.035)		(0.038)	
Other	186	6.785	255	6.699	0.087
		(0.058)		(0.048)	
Total	5,732	6.796	6,129	6.765	0.031**
		(0.010)		(0.009)	

Notes: Table 1B displays the number of observations, mean performance and standard error in brackets by country group and match type (with or without fans). "Fans" refers to all games played with fans in the stadium during period 2. "No Fans" refers to all games played without fans at the stadium during period 2. Period 2 refers to seasons 2020/21 and 2021/22. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4 Results

I run the same 10 regressions in both cases (Table 2A refers to the first period and Table 2B to the second one). These regressions are based on equation (1) from my empirical strategy to estimate the impact of fans on player performance. They control for $player \times team$ fixed effects to allow for the possibility that a player's productivity varies according to the team he plays for.[8] Furthermore, I also control for game-specific characteristics, an opponent-fixed effect and several weather characteristics. The first regression is straightforward and tests the impact of fans on the performance of all the players (column 1). Then I proceed by including each country group separately in the regression, which results in seven regressions since there are seven country groups (columns 2 to 8). In contrast, in the last two regressions, I include all country groups except English players, who serve as the base group (columns 9 and 10). Column 9 has $matchday$ fixed effects to control for conditions that may be particular to a specific part of the season and column 10 has $team \times matchday$ fixed effects to account for factors that are unique to an individual team and game, such as lineups.

In period 1 (Table 2A), I find that, on average, Premier League players experience no significant change in performance when supporters are gone (the coefficient of the variable $nofans$ in column 1 is negative but statistically insignificant). Looking at the country groups individually, I learn that some perform worse without fans, notably players from Africa, Latin America, South-eastern and Eastern Europe (excluding the EU) and the EU. For example, African players perform 2.6% worse in the post-lockdown period compared to the pre-lockdown period. A possible explanation is that these sub-groups may thrive off the energy provided by the fans meaning that fan engagement positively impacts their performance. One

[8] Results do not change when we include $player$ and $team$ fixed effects separately since only a small number of players change teams during the winter transfer window.

Table 2A: Effect of empty stadiums on players' performances in period 1

	(1)	(2) ENG	(3) AFR	(4) LAT	(5) SEE	(6) EU	(7) BI	(8) OTH	(9) All	(10) All
No Fans	-0.011	-0.017**	-0.010	-0.010	-0.011	-0.010	-0.010	-0.011	0.000	
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	
No Fans × England		0.018***								
		(0.006)								
No Fans × Africa			-0.016						-0.026**	-0.013
			(0.010)						(0.011)	(0.010)
No Fans × Latin				-0.011					-0.022**	-0.021**
				(0.010)					(0.011)	(0.010)
No Fans × S.E. + E. Europe					-0.018				-0.030*	-0.027*
					(0.015)				(0.016)	(0.015)
No Fans × EU						-0.004			-0.014**	-0.012*
						(0.006)			(0.007)	(0.006)
No Fans × B.I. (ex. Eng.)							-0.006		-0.017	-0.015*
							(0.010)		(0.010)	(0.009)
No Fans × Other								-0.008	-0.019	-0.023*
								(0.013)	(0.013)	(0.013)
Overall effect		**0.000**	**-0.026****	**-0.021***	**-0.029***	**-0.013***	**-0.016**	**-0.019**		
		(0.008)	**(0.012)**	**(0.012)**	**(0.017)**	**(0.008)**	**(0.011)**	**(0.013)**		
Match controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Player-team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Opponent team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Team-matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	7,475	7,475	7,475	7,475	7,475	7,475	7,475	7,475	7,475	7,459
R-squared	0.194	0.195	0.194	0.194	0.194	0.194	0.194	0.194	0.200	0.420

Notes: The regressand is the log of the players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. The base group in the last two columns is England. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 1 refers to season 2019/20. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

could also hypothesise that the players from these country groups may have struggled more mentally than players from the British Isles during and after the first lockdown due to UK travel restrictions. In the last two columns, I find that, in general, English players perform better without fans when compared to most other groups.

In period 2 (Table 2B), I find that, on average, Premier League players perform almost 1% better when fans are gone, which was not the case in the first period. In other words, players

Table 2B: Effect of empty stadiums on players' performances in period 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		ENG	AFR	LAT	SEE	EU	BI	OTH	All	All
No Fans	0.010**	0.007	0.010**	0.011**	0.010**	0.009*	0.011**	0.010**	0.029**	
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.014)	
No Fans ×		0.008**								
England		(0.004)								
No Fans ×			-0.008						-0.013	-0.007
Africa			(0.008)						(0.008)	(0.008)
No Fans ×				-0.010					-0.014*	-0.010
Latin				(0.007)					(0.007)	(0.007)
No Fans ×					-0.002				-0.007	-0.006
S.E. + E. Europe					(0.011)				(0.011)	(0.012)
No Fans ×						0.002			-0.003	-0.007
EU						(0.004)			(0.004)	(0.004)
No Fans ×							-0.011*		-0.015**	-0.013**
B.I. (ex. Eng.)							(0.006)		(0.007)	(0.006)
No Fans ×								-0.004	-0.009	-0.007
Other								(0.011)	(0.011)	(0.010)
Overall effect		**0.015*** **	**0.003**	**0.001**	**0.007**	**0.011****	**0.000**	**0.006**		
		(0.005)	**(0.008)**	**(0.008)**	**(0.011)**	**(0.005)**	**(0.007)**	**(0.011)**		
Match controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Player-team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Opponent team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Team-matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	11,861	11,861	11,861	11,861	11,861	11,861	11,861	11,861	11,861	11,846
R-squared	0.181	0.181	0.181	0.181	0.181	0.181	0.181	0.181	0.189	0.402

Notes: The regressand is the log of the players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. The base group in the last two columns is England. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 2 refers to seasons 2020/21 and 2021/22. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

performed worse when fans returned to the stadium. A more striking result is that English players attain ratings that are 1.5% higher without fans. In the same vein, players from the EU also appear to perform better without fans. For the remaining country groups, I do not detect any effect.

I repeat the same analysis for both cases but add one more restriction: whether the player played at home or away. In the first period, I found no effect of fans on the performance of players that played at home (all coefficients are statistically insignificant). The results can be found in Table 3A in the appendix. When I focus on players that played away (Table 3B in the appendix), we learn that some sub-groups perform worse without fans, notably players from Africa, Latin America, South-eastern and Eastern Europe (excluding the EU) and the British Isles (excluding England). In the last two columns of Table 3B (where English players act as the base group), on average, English players that play away perform better without fans than non-English players in away games without fans.

In the second period, on average, Premier League players attain 1.1% lower scores at home when there are no fans present (Table 3C). In addition, some country groups (players from Latin America, the British Isles (excluding England) and other players) perform worse at home without fans. Table 3D, which focuses on players that played away, displays the opposite result we found in Table 3C. On average, Premier League players that play away attain 1.4% higher ratings without fans than with fans in the stadium. Columns 2 to 8 reveal that, specifically, players from England and the EU perform better in away games without fans as compared to away games with fans. English players obtain ratings that are 2% higher in away games without fans than in away games with fans (column 2). In the last two columns of Table 3D, a reoccurring result is that English players who play away without fans perform better than the other country groups who play away without fans.

Table 3C: Effect of empty stadiums on players' performances in period 2, only home games

	(1)	(2) ENG	(3) AFR	(4) LAT	(5) SEE	(6) EU	(7) BI	(8) OTH	(9) All	(10) All
No Fans	-0.011**	-0.013**	-0.011**	-0.009*	-0.011**	-0.012**	-0.010**	-0.010**	-0.062	
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.039)	
No Fans × England		0.006								
		(0.005)								
No Fans × Africa			-0.002						-0.003	-0.001
			(0.011)						(0.015)	(0.014)
No Fans × Latin				-0.016					-0.000	-0.002
				(0.010)					(0.015)	(0.014)
No Fans × S.E. + E. Europe					0.001				-0.005	0.002
					(0.016)				(0.020)	(0.019)
No Fans × EU						0.005			-0.007	-0.005
						(0.006)			(0.010)	(0.009)
No Fans × B.I. (ex. Eng.)							-0.008		0.005	-0.001
							(0.009)		(0.015)	(0.013)
No Fans × Other								-0.013	-0.009	-0.010
								(0.014)	(0.020)	(0.019)
Overall effect		**-0.007**	**-0.012**	**-0.025****	**-0.009**	**-0.007**	**-0.018***	**-0.024***		
		(0.006)	**(0.011)**	**(0.011)**	**(0.016)**	**(0.006)**	**(0.010)**	**(0.014)**		
Match controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Player-team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Opponent team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Team-matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	5,973	5,973	5,973	5,973	5,973	5,973	5,973	5,973	3,739	3,732
R-squared	0.207	0.207	0.207	0.207	0.207	0.207	0.207	0.207	0.257	0.434

Notes: The regressand is the log of the players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. The base group in the last two columns is England. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 2 refers to seasons 2020/21 and 2021/22. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

5 Discussion and further tests

The previous results revealed that in the first period, specific non-English subgroups perform worse when games are played behind closed doors, especially when the focus is on away

Table 3D: Effect of empty stadiums on players' performances in period 2, only away games

	(1)	(2) ENG	(3) AFR	(4) LAT	(5) SEE	(6) EU	(7) BI	(8) OTH	(9) All	(10) All
No Fans	0.014***	0.010*	0.015***	0.014***	0.014***	0.014***	0.015***	0.014***	0.058	
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.041)	
No Fans × England		0.010*								
		(0.006)								
No Fans × Africa			-0.010						-0.039**	-0.021
			(0.011)						(0.016)	(0.014)
No Fans × Latin				-0.005					-0.043***	-0.040***
				(0.009)					(0.015)	(0.014)
No Fans × S.E. + E. Europe					-0.007				-0.058***	-0.060***
					(0.015)				(0.022)	(0.020)
No Fans × EU						0.000			-0.021**	-0.017**
						(0.006)			(0.010)	(0.008)
No Fans × B.I. (ex. Eng.)							-0.015*		-0.035**	-0.027**
							(0.008)		(0.014)	(0.011)
No Fans × Other								0.003	-0.027	-0.036**
								(0.016)	(0.017)	(0.017)
Overall effect		**0.020*** **	**0.005**	**0.009**	**0.006**	**0.014****	**0.000**	**0.017**		
		(0.006)	**(0.011)**	**(0.010)**	**(0.015)**	**(0.006)**	**(0.009)**	**(0.016)**		
Match controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Player-team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Opponent team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Team-matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	5,882	5,882	5,882	5,882	5,882	5,882	5,882	5,882	3,732	3,732
R-squared	0.209	0.209	0.209	0.209	0.209	0.209	0.209	0.209	0.256	0.256

Notes: The regressand is the log of the players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. The base group in the last two columns is England. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 2 refers to seasons 2020/21 and 2021/22. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

games. This result may come as a surprise since one could hypothesise that players perform better when there are no fans in away games. In other words, when games are played behind closed doors, the home advantage is reduced, increasing the away team's both offensive and defensive efficiency (Destefanis et al. 2022). One could argue that the players may have needed time to adjust to the new environment in away games, which could explain the initial drop in

players' performance in those ghost games. Period 2 results differ largely from period 1: on average, Premier League players perform slightly better without fans, especially English and EU players (Table 2B). Upon closer inspection, this only holds true in away games without fans (Table 3D), which is consistent with the recent literature (Destefanis et al. 2022; Leitner & Richlan 2021; McCarrick et al. 2021). Similarly, players perform worse on average when they play at home without fans, which could be proof of the decreased home advantage. However, this result does not apply to all country groups.

In the remaining part of this section, I will focus on English players and their performance in period 2. English players perform 2% better in away games without fans than in away games with fans (Table 3D). The coefficient is statistically significant at all levels. When compared to the other groups, they also perform significantly better. The goal is to explain why English players attain higher scores without fans by documenting a series of tests.

5.1 Youth team, place of birth and experience

One can argue that a player who, at youth level played for the same team he currently plays for, is more likely to be attached to that team than a player who did not. It seems reasonable to assume that such a player would understand more about the club's history and the importance of certain games. For this reason, it could be that those types of players perform better in away games without fans since there is no pressure from the home team fans that could negatively affect their performance. To test this possibility, I introduce a new dummy variable in our equation (equal to 1 if the player played for the team at the youth level and 0 otherwise). The dummy is included as an interaction with the $nofans$ variable (Table 4, column 1).[9] Contrary to my expectations, I find the opposite result. English players, who played at the youth level for

[9] In Table 4, I drop all observations of non-English players, and I only focus on away games from the second scenario.

the same team they currently play, perform on average 1.8% worse in away games without fans than those who did not play for the same team at the youth level. Perhaps, former academy players of their team experience more motivational hurt as they may strive on the energy of the opposition's fans.

In the same manner as before, the place of birth could be another indicator of a player's attachment to the club. A player born close to his current team's stadium is more susceptible to supporting the team than someone born far away. To test this hypothesis, I collect information on all English players' places of birth. I use the Haversine formula to determine the distance between a player's place of birth and his current team's stadium.[10] I proceed by creating a new dummy based on the distances I calculated before (equal to 1 if the player was born less than 45 kilometres away from the stadium and 0 otherwise) to decide whether a player was born near the stadium or not. As before, I include this dummy as an interaction with the $nofans$ variable (Table 4, column 2). However, I do not find any evidence that English players born near their current club's stadium perform better in away games without fans than those who were not born near the stadium (the coefficient of the interaction term $nofans \times place\ of\ birth$ is not statistically significant).

The following hypothesis I want to test is whether less experienced English players performed better in away games without fans than more experienced players. Experience is referred to as the time a player has been at the club for.[11] I create a new dummy $experience$ (equal to 1 if the player has more than two years of experience at the club and 0 otherwise), which I include in the equation as an interaction term with the $nofans$ variable (Table 4, column 3). The coefficient of the interaction term is insignificant, meaning that we cannot

[10] The Haversine formula calculates the great-circle distance between two points on a sphere based on their longitudes and latitudes.
[11] Since some players have already previously played for the club, I also include that previous experience in my calculation.

Table 4: Effect of empty stadiums on English players' performances in period 2 (only away games), alternative mechanisms

	(1) Youth team	(2) Place of birth	(3) Experience
No Fans	0.026***	0.024***	0.017**
	(0.007)	(0.007)	(0.008)
No Fans × Youth Team	-0.018*		
	(0.010)		
No Fans × Place of Birth		-0.009	
		(0.009)	
No Fans × Experience			0.006
			(0.009)
Match controls	✓	✓	✓
Player-team FE	✓	✓	✓
Opponent team FE	✓	✓	✓
Observations	2,247	2,247	2,247
R-squared	0.229	0.228	0.228

Notes: The regressand is the log of English players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. Youth Team is a dummy equal to 1 if the player played for the team at the youth level and 0 otherwise. Place of Birth is a dummy equal to 1 if the player was born less than 45 kilometres away from the stadium and 0 otherwise. Experience is a dummy equal to 1 if the player has more than two years of experience at the club and 0 otherwise. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 2 refers to seasons 2020/21 and 2021/22. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

conclude that less experienced players of English nationality perform better in away games without fans than English players with more experience.

5 Conclusion

Using Premier League data, the highest professional football league in England, I investigate the effects of fan presence in stadiums on the performance of different player groups. I use three seasons (2019/20, 2020/21, and 2021/22) and split them into two periods. Period 1 refers to season 2019/20, played with fans until the first lockdown in March 2020 and without fans until the end of the season. Period 2 refers to seasons 2020/21, played mainly without fans, and 2021/22, played entirely with fans. I find that in period 1, almost all player groups not from the

British Isles perform worse without fans in the stadium. For instance, African players are one of the groups that perform worse without fans during that period. This is the opposite result found in a similar study which focused on the Serie A (Caselli, Falco & Mattera 2021). In period 2, I find that, in general, players perform worse in home games without fans than in home games with fans, while the opposite result holds in away games. This result is consistent with the recent findings of reduced home advantage in the absence of fans (Destefanis et al. 2022; Leitner & Richlan 2021; McCarrick et al. 2021). In particular, English players perform better in away games without fans as compared to away games with fans. In my attempts to find an explanation for this result, I conclude that specifically English players that are not former academy players of their team perform better in away games without fans than English players who played for their current team at the youth level.

These findings are striking because they reveal that even the best athletes can be affected by the lack of fans, whether this leads to better or worse performances. Though it may seem surprising, players at the top of their game are likely to be fazed by the absence of fans as it may relieve performance anxiety or hurt their motivation. Furthermore, it is important to note that the results do not remain consistent throughout the two periods. A possible explanation could be that we went from having fans to having no fans in the first period, whereas we had the opposite case in the second. Understanding why athletes react differently to these two cases could be another interesting topic for future research.

Finally, my findings indicate that there can be other financial repercussions caused by the COVID-19 pandemic apart from the most obvious one, which is the matchday revenue lost due to games being played behind closed doors. Supposing that the absence of fans negatively affects the performance of football players (as was the case for most player groups in period 1) and teams, could further lead to more economic harm because an underperforming player is less likely to be offered a new contract. In the same vein, underperforming teams are less likely

to participate in or even win specific competitions that generate substantial amounts of money for the club.

I acknowledge that this study has some limitations, such as the possibility that there were variables I was unable to control for that could have affected my results, such as changes in teams' training routines or the mental impact of the first lockdown on players. However, I hope to have contributed to the contemporary literature on the effects of the COVID-19 pandemic on football.

References

BBC. 2022. "Granit Xhaka: Arsenal midfielder came close to quitting club over 'pure hate' from supporters." BBC Sport. April 13, 2022.
https://www.bbc.com/sport/football/61093079

BeSoccer. 2020. "How important are match-going football fans?" Accessed December 12, 2022.
https://www.besoccer.com/new/how-important-are-match-going-football-fans-844739

Bransen, Lotte. 2019. "Choke or shine? Quantifying football players' abilities to perform under mental pressure." SciSports. February 27, 2019.
https://www.scisports.com/choke-or-shine-quantifying-football-players-abilities-to-perform-under-mental-pressure/#

Cable, Jonathan, Daniel Kilvington, and Glyn Mottershead. 2022. "'Racist behaviour is interfering with the game': exploring football fans' online responses to accusations of racism in football." *Soccer & Society* 23 (8): 880-893.
https://doi.org/10.1080/14660970.2022.2109804

Caselli, Mauro, Paolo Falco, and Gianpiero Mattera. 2021. "When the Stadium Goes Silent: How Crowds Affect the Performance of Discriminated Groups." DEM Working Papers N. 2021/1.
https://doi.org/10.1086/719967

Cruyff, Johan. 2016. *My Turn: The Autobiography.* United Kingdom: Pan MacMillan

Destefanis, Sergio, Francesco Addesa, and Giambattista Rossi. 2022. "The impact of COVID-19 on home advantage: a conditional order-m analysis of football clubs' efficiency in the top-5 European leagues." *Applied Economics* 54 (May): 6639-6655. https://doi.org/10.1080/00036846.2022.2074361

Kausel, Edgar E., Santiago Ventura, and Arturo Rodríguez. 2019. "Outcome bias in subjective ratings of performance: Evidence from the (football) field." *Journal of Economic Psychology* 75 Part B (December): 102132. https://doi.org/10.1016/j.joep.2018.12.006

Leitner, Michael C., and Fabio Richlan. 2021. "No more advantage for the home team: Refereeing without social pressure in European elite football during the COVID-19 pandemic." *Frontiers in Sports and Active Living* 3 (September). https://doi.org/10.3389/fspor.2021.720488

McCarrick, Dane, Merim Bilalic, Nick Neave, and Sandy Wolfson. 2021. "Home advantage during the COVID-19 pandemic: Analyses of European football leagues." *Psychology of Sport and Exercise* 56 (September): 102013. https://doi.org/10.1016/j.psychsport.2021.102013

McMurray, Calli. 2021. "How Playing in Empty Stadiums Affects Athletes." BrainFacts.org. July 26, 2021.
https://www.brainfacts.org/thinking-sensing-and-behaving/thinking-and-awareness/2021/how-playing-in-empty-stadiums-affects-athletes-072621

Principe, Francesco and Jan C. van Ours. 2022. "Racial bias in newspaper ratings of professional football players." *European Economic Review* 141 (January): 103980.
https://doi.org/10.1016/j.euroecorev.2021.103980

Premier League. 2020. "How has the COVID-19 pandemic affected Premier League matches?" June 15, 2020.
https://www.premierleague.com/news/1682374

Psychology Educator. 2013. "The Psychology of Sports Fans: How Fans Affect Players." Accessed December 12, 2022.
https://psychologyeducator.wordpress.com/2013/08/15/the-psychology-of-sports-fans-how-fans-affect-players/

We Make Footballers. 2022. "Why are football fans important?" May 26, 2022.
https://www.wemakefootballers.com/news/why-are-football-fans-important

WhoScored.com. "WhoScored Ratings Explained." Accessed December 12, 2022.
https://www.whoscored.com/Explanations

Appendix

Table 3A: Effect of empty stadiums on players' performances in period 1, only home games

	(1)	(2) ENG	(3) AFR	(4) LAT	(5) SEE	(6) EU	(7) BI	(8) OTH	(9) All	(10) All
No Fans	-0.004	-0.005	-0.003	-0.004	-0.003	-0.002	-0.004	-0.003	-0.062	
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.039)	
No Fans ×		0.004								
England		(0.008)								
No Fans ×			-0.003						-0.003	-0.001
Africa			(0.013)						(0.015)	(0.014)
No Fans ×				0.004					-0.000	-0.002
Latin				(0.014)					(0.015)	(0.014)
No Fans ×					-0.002				-0.005	0.002
S.E. + E. Europe					(0.020)				(0.020)	(0.019)
No Fans ×						-0.006			-0.007	-0.005
EU						(0.008)			(0.010)	(0.009)
No Fans ×							0.008		0.005	-0.001
B.I. (ex. Eng.)							(0.015)		(0.015)	(0.013)
No Fans ×								-0.008	-0.009	-0.010
Other								(0.020)	(0.020)	(0.019)
Overall effect		**-0.001**	**-0.007**	**0.000**	**-0.006**	**-0.008**	**0.004**	**-0.011**		
		(0.009)	**(0.014)**	**(0.015)**	**(0.021)**	**(0.010)**	**(0.017)**	**(0.020)**		
Match controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Player-team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Opponent team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Team-matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	3,739	3,739	3,739	3,739	3,739	3,739	3,739	3,739	3,739	3,739
R-squared	0.235	0.235	0.235	0.235	0.235	0.235	0.235	0.235	0.257	0.434

Notes: The regressand is the log of the players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. The base group in the last two columns is England. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 1 refers to season 2019/20. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 3B: Effect of empty stadiums on players' performances in period 1, only away games

	(1)	(2) ENG	(3) AFR	(4) LAT	(5) SEE	(6) EU	(7) BI	(8) OTH	(9) All	(10) All
No Fans	-0.011	-0.021***	-0.009	-0.008	-0.010	-0.010	-0.008	-0.010	0.058	
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.041)	
No Fans × England		0.030***								
		(0.009)								
No Fans × Africa			-0.022						-0.039**	-0.021
			(0.015)						(0.016)	(0.014)
No Fans × Latin				-0.027**					-0.043***	-0.040***
				(0.013)					(0.015)	(0.014)
No Fans × S.E. + E. Europe					-0.038*				-0.058***	-0.060***
					(0.021)				(0.022)	(0.020)
No Fans × EU						-0.000			-0.021**	-0.017**
						(0.008)			(0.010)	(0.008)
No Fans × B.I. (ex. Eng.)							-0.019		-0.035**	-0.027**
							(0.013)		(0.014)	(0.011)
No Fans × Other								-0.011	-0.027	-0.036**
								(0.016)	(0.017)	(0.017)
Overall effect		**0.008**	**-0.031****	**-0.035****	**-0.048****	**-0.011**	**-0.028****	**-0.021**		
		(0.010)	**(0.016)**	**(0.014)**	**(0.022)**	**(0.009)**	**(0.014)**	**(0.017)**		
Match controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Player-team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Opponent team FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
Team-matchday FE	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	3,732	3,732	3,732	3,732	3,732	3,732	3,732	3,732	3,732	3,732
R-squared	0.233	0.236	0.233	0.234	0.233	0.233	0.233	0.233	0.256	0.452

Notes: The regressand is the log of the players' ratings. No Fans is a dummy variable equal to 1 if the game was played without fans. The base group in the last two columns is England. Match controls include home dummy, derby dummy, home dummy interacted with derby dummy and No Fans and weather conditions. Period 1 refers to season 2019/20. Standard errors (in parentheses) are clustered at the team-matchday level.
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$