ISSN: 1756-8811

Birkbeck Sport Business Centre

Research Paper Series

Football clubs' efficiency and COVID-19 in the Big-5 European leagues

Sergio Destefanis

University of Salerno

Francesco Addesa

Leeds Beckett University

Giambattista Rossi

Birkbeck University of London

Volume 13, Number 2, September 2021

Abstract

This study aims to contribute to the recent literature on the effects of COVID on football teams' performance, focusing on the impact of playing behind closed doors – due to the health and safety measures following the COVID-19 pandemic outbreak - on offensive and defensive technical efficiency. Using a long season-level dataset for the top 5 European leagues, a novelty for efficiency studies on football, the analysis compares the ten seasons (2009-10 to 2018-19) played before the pandemic outbreak with the only season (2020-21) entirely played behind closed doors. The methodology applied to calculate the efficiency scores is the conditional order-m, whose application represents a further novel contribution to the literature on football teams' efficiency. Our findings are consistent with the recent literature on the impact of ghost games on teams' performance and show an erosion of the home advantage, likely due to the reduced pressure on visiting teams deriving from the lack of home crowd support.

Keywords: efficiency, conditional order-m, ghost games, home advantage, football, COVID-19

1. Introduction

In every competitive industry, firms and organisations constantly try to efficiently use their resources. This implies that, given specific amounts of factor inputs, an organisation, or a decision-making unit (DMU) must generate the maximum potential outputs or utilise the minimum necessary inputs, considering constraints such as physical quantities and technical relationships (Coelli et al., 2005). The football industry also follows this line of conduct as football clubs constantly analyse and appraise their performance on-and-off the pitch in different managerial scenarios, from buying football players to signing commercial deals. In doing so, the football industry, just as any other industry, has come to realise that the data it produces represent a critical resource for the monitoring, management, and analysis of its own behaviour. The emergence of new modes of data collection, more granular than those previously available, has increased this awareness.

In this article, we rely on data routinely produced and stored on the website www.whoscored.com to estimate football clubs' on-field performance efficiency and evaluate the impact on this efficiency of the so-called ghost matches policy brought about by the COVID-19 pandemic. We construct and use a panel dataset over a period of 12 seasons from 2009-10 to 2020-21 for the top 5 European leagues: English Premier League, Spanish La Liga, Italian Serie A, German Bundesliga and French Ligue 1, contributing to the operational research literature in sport in three main aspects. First of all, we jointly consider the top 5 European leagues throughout a long time period and assess the role that some contextual factors may have on the offensive and defensive efficiency of their teams. Then, we include among these factors the COVID-19 pandemic and appraise the erosion of the home advantage that has followed its outbreak. Finally, we adopt the conditional order-m approach (Daraio and Simar, 2005; De Witte and Kortelainen, 2013), that provides more robust and reliable estimates of efficiency and of the role of contextual factors presiding to its determination.

Academic research on football efficiency has been quite extensive. Previous articles have analysed football clubs' efficiency, in financial and/or sporting terms, in major and minor national championships and knockout competitions: the English Premier League (Dawson et al., 2000; Haas, 2003; Barros and Leach, 2006); the Spanish Liga (Espitia-Escuer and Garcia-Cerbian, 2004; Gonzalez-Gomez and Picazo-Tadeo, 2010; Barros and Garcia-del-Barrio, 2011); the Italian Serie A (Boscà et al., 2009; Rossi et al., 2019); the German Bundesliga (Tiedermann et al., 2011); the Portuguese Primeira Liga (Ribeiro and Lima, 2012); the Greek Super League 1 (Barros and Douvis, 2009); the Brasilian Serie A (Barros et al., 2010); the Mexican Liga MX (Davila and Garcia-Cebrian, 2012); the UEFA Champions League (Espitia-Escuer and Garcia-Cerbian, 2010; Zambon-Ferraresi et al., 2017).

In their comprehensive literature review, Kulikova and Goshunova (2013) highlight that a non-parametric frontier methodology, the Data Envelopment Analysis (DEA), is the most common methodological approach to analyse technical efficiency in sport. Furthermore, several studies investigate football clubs' on-field efficiency following a two-stage analysis format where efficiency measurement is carried out through DEA and the evaluation of efficiency determination is conducted through regression analysis on the scores obtained at the previous stage (examples of this procedure also include such important studies as Garcia-Sanchez, 2007; Boscà et al., 2009; Sala-Garrido et al., 2009; Villa and Lozano, 2016). However, this technique suffers from well-known problems of validity and inference. Firstly, the DEA estimator has, under many circumstances, less than root-n convergence to the true production frontier (Simar and Wilson, 2015). Secondly, being a full frontier estimator and enveloping all the points in the production set, DEA is sensitive to outliers (Kneip et al., 1998). Finally, DEA does not straightforwardly allow for contextual variables - variables potentially relevant for the production process but beyond the control of the productive units - to influence the support of the production process. Having considered all these issues, we adopted an alternative and more recently developed procedure, the conditional order-m (Daraio and Simar, 2005), which builds upon the order-m approach proposed in Cazals et al. (2002). Being a partial frontier estimator, the order-m technique is more robust than other non-parametric estimators. Moreover, it routinely has root-n convergence to the true production frontier (Simar and Wilson, 2015). What is more, for our purposes conditional order-m, and in particular its extension developed in De Witte and Kortelainen (2013), consistently estimates the impact that a discrete event, such as the introduction of ghost matches, may have on the determination of efficiency. To the best of our knowledge, we apply for the first time conditional order-m to the analysis of on-field efficiency of football teams.

While Neale (1964) and Szymanski (2003) argue that, since football clubs are grouped into leagues, on-field efficiency should be calculated considering their respective leagues as DMUs, Sloane (1971) conversely argues that the unit of analysis for performance should be the club itself. In our context, we select the latter approach for the case of clubs competing in the top 5 European leagues throughout a long period. We stress that this is by itself a further contribution to the literature, as previous research has never considered the measurement of efficiency in the context of more than two leagues (Boscà et al., 2009).

Yet, the contributions that we have highlighted above with respect to empirical method and sample must be considered in the light of the focus of this article. Our analysis investigates the reduction of the home advantage factor during the 2020-21 football season. The outbreak of the COVID-19 pandemic in 2020 allows us to investigate how the home advantage, one of the most studied and best documented phenomena in sports (Courneya and Carron, 1992; Pollard and Pollard, 2005), was eroded in the absence of spectators. We use this natural experiment scenario to assess how professional football clubs in the top 5 European leagues have reacted to this peculiar circumstance, particularly insofar as their offensive and defensive efficiency was concerned.

The article is organised as follows. In Section 2, we present a review of the literature on home advantage, focusing especially on the most recent contributions following the pandemic outbreak. In Section 3 we describe the methodology adopted for the empirical analysis, whereas Section 4 shows and discusses our main results. Section 5 concludes.

2. Literature review

The phenomenon of home advantage is defined as "*the consistent finding that the home teams in sport competition win over 50% of the games played under a balance home and away schedule*" (Courneya and Carron, 1992, p. 13). Stadium spectators are mostly supporters of local home teams, whose winning probability increases due to their fans' support from the stands (Goumas, 2014; Nevill et al., 1996; Pollard, 2006; Pollard and Gomez, 2009; Ponzo and Scoppa, 2018). Despite being a worldwide phenomenon, with variation both over time and across regions, overwhelming empirical evidence (Pollard, 1986; Jamieson, 2010; Legaz-Arrese et al., 2013) shows that home advantage is more prevalent in team sports, such as football in particular, than individual sports.

To explain this phenomenon, researchers have focused their attention towards several mechanisms based on the relative importance and interplay of three main factors: home crowd support affecting visiting teams' performance and referee's decisions (Schwartz and Barksy, 1977; Garicano et al., 2005; Pollard and Pollard, 2005; Buraimo et al., 2010); home teams' familiarity with local playing conditions (Pollard, 1986; Barnett and Hilditch, 1993; Clark and Norman, 1995; Moore and Brylinsky, 1995); visiting teams' travel fatigue (Courneya and Corran, 1992; Ponzo and Scoppa, 2018; Van Damme and Baert, 2019). Disentangling the influence of these factors requires the experimental manipulation of real-world sport events, but this condition is almost impossible to meet. Alternatively, evidence can primarily be obtained on rare cases with limited sample sizes either indirectly, through the analysis of matches, teams or dimensions with varying attendance and travel burden by drawing conclusions from the characteristics of countries with a varying degree of home advantage, or more directly by considering special circumstances, such as same-stadium derbies, teams moving to a new sport facility or city, or spectators' ban due to hooligan violence.

The outbreak of COVID-19 pandemic in the early months of 2020 has offered a unique opportunity to carry out a virtual experiment to assess whether matches played in complete absence of spectators - known as ghost matches - significantly alter the home advantage effect. Whilst two out of the three main factors explaining home advantage - home teams' familiarity with local conditions and visiting teams' travel fatigue - have not been affected by the pandemic outbreak, the health and safety measures imposing behind-closed-door games have removed the third factor – the crowd pressure on away teams and referees. Indeed, the most recent studies – as shown in Table 1 - suggest that the COVID pandemic significantly affected home advantage. Using data from the German Bundesliga's 2019-2020 season, Reade and Singleton (2020) find that the home advantage is eroded in the first ghost games to be then partially recovered at the end of the season, when players became accustomed to empty stadia. Home teams won just 32% of ghost matches compared with 43% in the same season before the pandemic outbreak, while away teams won 45% of the ghost games, compared with 35% of games played with spectators on the stands. Further German evidence comes from Tilp and Thaller (2020) and Fischer and Haucap (2021), who find that the ghost games' impact on home advantage – measured by the game outcomes (home win, draw or away win) – is very small in lower divisions as teams mostly play in half-empty stadia. In other words, only teams that are used to playing in a full stadium are likely to be affected by ghost games. Correia-Oliveira and Andrade-Souza (2021) analyse data from the first and second divisions of the English, German, Italian and Spanish leagues from 2016/17 to 2019/20. They also find evidence that home advantage – in this case measured by the ratio between the number of points won at home and the total number of points won at home and away - is reduced in games played after the outbreak of the COVID-19 pandemic. Ferraresi and Gucciardi (2021) find similar results in relation to the penalty kicks, as the probability of missing a penalty increases for home teams and decreases for away teams in the ghost games played in the top 5 European leagues.

Table 1 about here

Partially different results are found by Bryson et al. (2021) on a wider dataset of 23 professional leagues from 17 countries. These authors find that ghost games have no effect on the likelihood of a home win, the goal difference and the total goals scored. However, they find that ghost games have a negative impact on the home advantage in terms of yellow cards. This in accordance with both Endrich and Gesche (2020), McCarrick et al. (2020), and Reade et al. (2020), who find evidence that referees' bias is reduced in ghost games as the relationship between referees, players and coaches is rebalanced in the absence of social pressure from spectators. Also Scoppa (2021) – using data from first and second divisions of five European Leagues (England, Germany, Italy, Portugal and Spain) - finds consistent evidence that the home advantage, measured by various performance indicators (points, goals, shots, etc.), is almost halved in ghost matches, and referees' decisions are much more balanced without the crowd pressure. Inconsistent evidence is instead provided by Ramchandani and Miller (2021), who analyse the first division of five European Leagues (England, Germany, Italy, Portugal and Spain) and – using the same measure as in Correia-Oliveira and Andrade-Souza (2021) - find that ghost games have a negative impact on home advantage only in Germany and Italy; and by Benz and Lopez (2021), who focus on goal and yellow card difference in 17 European leagues and show a negative impact of ghost games on home advantage in some leagues and a positive impact in others.

In all these papers, the natural experiment opportunity given by the COVID-19 pandemic was exploited to provide clear evidence about how football clubs' performance was affected by playing behind closed doors. However, the metrics used was limited to few key variables, and never included the measurement of technical efficiency. In this paper we fill this gap in the literature by considering an extensive production set, based on data from the top 5 European leagues (England, France, Germany, Italy and Spain) for the last 12 seasons. We provide measures of football clubs' efficiencies before and after the onset of ghost matches and, modelling this event as a contextual variable, we give a quantitative assessment of its impact on the technical efficiency of football clubs.

3. Conditional Order-M and Regression

A longstanding issue in the field of nonparametric efficiency analysis relates to the treatment of contextual variables, that is, variables that are potentially relevant for the production process but also beyond the control of the productive units under observation. A first traditional solution to this problem is based on the partition of the sample according to different categories of contextual variables and is only applicable when the contextual variable is binary or categorical and is likely to involve a considerable reduction of the sample over which efficiency scores are computed, and the direction of the influence of the contextual variable on efficiency is not known a priori (Charnes et al., 1981). Another solution involves the comparison of the observations under scrutiny with other observations that have the same or a more detrimental value of the contextual variable (Banker and Morey, 1986). In this case, the influence of the contextual variable on efficiency must be known a priori, and there may be computational problems of the contextual variable being not continuous. Hence, the above solutions imply that only either continuous or discrete contextual variables can be considered and make it very difficult to deal with several contextual variables simultaneously. These strictures have prompted many researchers to resort to a two-stage approach, involving first the calculation of the efficiency scores without any allowance for the contextual factors, and then the appraisal of the role of these factors through regression analysis. There are various issues with this approach, the main one arguably being the requirement for the so-called separability condition to be fulfilled, as pointed out by Simar and Wilson (2007). This condition effectively requires that contextual variables do not affect the shape of the attainable input–output set and can hardly be thought to be fulfilled a priori in many situations, including the present one. Why production sets should be in principle uniform across countries, or uniformly touched by post-pandemics institutional changes? Hence, in this article we do not rely on the assumption of separability among inputs, outputs, and contextual variables. Accordingly, for our analysis we adopt the conditional order-m approach, initiated by Cazals et al. (2002), Daraio and Simar (2005) and further developed by De Witte and Kortelainen (2013)[1].

The conditional order-m approach is based on the fundamental idea of using contextual variables to identify the most similar observations, and estimating the efficiency around windows of these similar observations. Comparing the efficiency scores obtained unconditionally and conditionally on this similarity yields information about the impact of contextual variables that does not require the separability assumption, does not require any assumption on the direction of the influence of contextual variables on efficiency and allows for many contextual variables - both continuous and discrete - to be brought into play at the same time.

More formally, for a given input vector $X \in \Re_{+}^{p}$, used to produce output vector $Y \in \Re_{+}^{q}$, (unconditional) order-m output-oriented efficiency will be estimated by:

[1] In principle, we could test for the existence of the separability condition through the test suggested in Daraio et al. (2018). This test, however, runs into very serious computational problems, especially when several contextual variables must be considered at the same time (Wilson, 2020, p. 60). Besides, it is not likely that the separability assumption could hold in our case, as was already noticed above.

$$(1)\quad \hat{\theta}_{m,n}(x,y) = \int_0^\infty [1 - (1 - \breve{F}_{Y|X,n}(uy|x))^m]du$$

where $\breve{F}_{Y|X}$ is the empirical counterpart of the survival function $F_{Y|X}(y|x)$ defined over a subset of m observations randomly drawn with replacement from a set of n observations. The partial frontier obtained in this way is less sensitive to outliers than the global frontier comprising all n observations. Observations with an efficiency score above unity are inefficient (their outputs must be expanded in a proportion of $(\hat{\theta}_m - 1)$ to reach the order-m frontier). Observations with an efficiency score equal to unity are efficient (they are located on the order-m frontier). Finally, observations with an efficiency score below unity are classified as 'super-efficient'. In empirical applications, the size of the m subset is chosen in a way that stabilizes the proportion of super-efficient observations in the sample. As we have already noted, the order-m approach is more robust to outliers than most non-parametric estimators, being a partial frontier estimator. Moreover, this estimator routinely has root-n convergence to the true production frontier (Simar and Wilson, 2015).

Next, for a given input vector $X \in \mathfrak{R}_+^p$, used to produce output vector $Y \in \mathfrak{R}_+^q$, in an environment characterized by contextual variables $Z \in \mathfrak{R}_+^r$, estimation of conditional order-m output-oriented efficiency will be given by:

$$(2)\quad \hat{\theta}_{m,n}(x,y|z) = \int_0^\infty [1 - (1 - \breve{F}_{Y|X,Z,n}(uy|x,z))^m]du$$

Now, $\breve{F}_{Y|X,Z,n}$, the empirical counterpart of the survival function $F_{Y|X,Z}(y|x,z)$ is defined over a subset of m observations drawn from a set of n observations, *with the probability of drawing a given subset depending on a value of z and the kernel function estimated around this value*. Note the difference with the unconditional case, where the probability of being drawn is equal for all observations. As a result, in the conditional model one compares like with like and does not need to impose any assumption of separability on the data generating process. In the present application, the definition of the kernel function allows for continuous, ordered discrete (categorical), and unordered discrete (binary) variables, along the lines of De Witte and Kortelainen (2013). The kernel bandwidths for these variables are estimated through least-squares cross validation (Li and Racine, 2008).

Once estimates of both unconditional and conditional order-m efficiency are available, one can compute for each observation the ratio:

$$(3) \quad \frac{\hat{\theta}_{m,n}(x, y|z)}{\hat{\theta}_{m,n}(x, y)}$$

Regressing (3) over the Z variables yields estimates of the influence of these variables on efficiency that are not liable to the criticisms levelled to the traditional two-stage approach (see on this De Witte and Kortelainen, 2013, p. 2406). In order to do so, we rely on nonparametric regression analysis. This flexible specification does not impose any functional form on the relationship between efficiency scores and the contextual variables, allowing at the same time the calculation of standard errors for the regression coefficients (Li and Racine, 2007).

4. Empirical analysis and results

4.1. Production set and contextual variables

Our analysis covers 12 seasons (2009-10 to 2020-21) of the top 5 European domestic leagues according to the Union of European Football Associations (UEFA) ranking: English Premier League, Spanish La Liga, Italian Serie A, German Bundesliga and French Ligue 1. Since we rely on season-level team data drawn from the website www.whoscored.com, the 2019-20 season was excluded from the analysis as it is not possible to differentiate between pre-COVID and post-COVID production sets. Moreover, unlike the other four leagues under investigation, French Ligue 1 did not resume and complete the aforementioned season after the suspension due to the pandemic outbreak. Therefore, our dataset comprises a total of 1078 observations: 980 pre-COVID and 98 post-COVID.

Our production sets are differentiated by offense and defence, home and away games, and are broadly based on Boscà et al. (2009)[2]. Our contextual variables include, of course, a post-COVID binary variable, and other factors that we describe below. These variables, as well as the variables included in our production sets are listed in Table 2, whereas Table 3 presents the main descriptive statistics. Since our main aim is to assess the impact of the absence of crowd due to the health and safety measures implemented after COVID outbreak on teams' offensive and defensive efficiency, we present these statistics differentiating between the pre-COVID and post-COVID situations. It clearly emerges that, on the one hand, teams' home performance has worsened in relation to all the three output variables considered: points gained (average decreased from 1.64 to 1.45), goals scored (average decreased from 1.55 to 1.48) and goals conceded (average increased from 1.16 to 1.32). On the other hand, teams' away performance has significantly improved, with more points gained (from 1.11 to 1.29) and goals scored (from 1.16 to 1.32) and fewer goals conceded (from 1.55 to 1.48). It is not clear how these changes can relate to changes in either inputs or outputs. Possession does not change significantly, and we notice a generalised decrease in home and away shots, through balls and crosses, both made and conceded (this reduction is particularly significant for through balls and less strong for crosses). Tackles also decrease, whereas there is an increase in dribbles. These two developments also apply, although with varying strength, both to home and away games. Overall, the hypothesis that there may be variations in offensive and defensive technical efficiency seems to be worth inquiring.

[2] In the baseline production set, balls kicked into the opposing team's area are proxied by through balls. We proxy the same variable by crosses in a production set used for a robustness check.

Table 2 and 3 about here

Paramount among our contextual variables is a post-COVID binary variable that requires no further comment here. Other contextual variables potentially impacting on teams' efficiency are a) the number of managers employed during a season, that is used as a proxy for an unlucky season (teams that incur an unlucky season, possibly due to random events, change manager once or more than once; we do not ascribe any causal influence to the managers' change), and 2) the number of penalties attempted for the offensive efficiency, the number of penalties converted by the opponents for the defensive efficiency[3]. The penalties variables include zero values. Therefore, it is difficult to use them as inputs, while it is, in our opinion, appropriate and novel to use them as determinants of efficiency.

Like possession, the number of managers employed does not appreciably change before and after the COVID-19 outbreak. On the other hand, the number of penalties converted by the opponents at home games, and the number of penalties attempted at away games, significantly increase. One can relate these findings to similar results in the literature for other variables related to the referees' behaviour. We will illustrate below their connection to the changes that we have noticed in the offensive and defensive outputs.

The contextual variables also include binary variables for all leagues (the English Premier League being the default category), a placebo binary variable for 2018-19 and a set of binary variables for the introduction of the Video Assistant Referee (VAR). The inclusion of the league dummies is by itself a novelty for the literature and will be commented again below. The placebo variable is a safeguard against our results being driven by unknown factors evolving over time independently of the COVID-19 outbreak (when including this variable,

[3] The difference in the penalty variables used for offensive and defensive efficiency is due to the lack of data for the number of penalties attempted by the opponents in our dataset. However, our data show high correlation between penalties attempted and penalties scored, therefore our results are not heavily affected.

estimation was carried out from 2010-11 to 2018-19). The introduction of the VAR was modelled through a series of binary variables appropriately differentiated across countries, as VAR was introduced in 2017-18 in Italy and Germany, in 2018-19 in Spain and France and only in 2019-20 in England.

4.2. Findings

The main descriptive statistics of the efficiency scores for both unconditional and conditional order-m are presented in Tables 4 and 5, respectively for offensive and defensive efficiency[4]. These results are obtained for the baseline production set, where balls kicked into the opposing team's area are proxied by through balls. In the Appendix we show the results from a robustness check of our specification done by replacing through balls with crosses. Results are qualitatively unchanged, although efficiency is slightly lower (therefore we chose the set with through balls as our baseline case). We provide output-oriented scores that should vary in principle between zero and infinity. In the case of order-m, we have however super-efficient scores with a score lower than one. The following remarks are in order. First, the technical efficiency measures increase as contextual variables are included in the analysis. This is particularly true when penalties are included along with league dummies, number of managers and the post-COVID dummy. When carrying out these calculations, league dummies and the post-COVID dummy were, of course, treated as binary variables, while the number of managers was treated as a categorical variable.

Tables 4-5 about here

Second, home efficiency is in principle higher than away efficiency, both for offense and defence. Yet, including contextual variables marginally reverses this ordering, at least for

[4] These scores were computed through a script written in R. We gratefully acknowledge Kristof De Witte for his generosity in sharing his scripts with us.

defensive efficiency. A final remark is that mean and median efficiencies are high, especially if allowance is made for contextual variables. Given the respectable size of our sample (1078 observations), this vouchsafes for a good specification of our augmented production set.

Regressing the appropriate version of (3), with and without penalties, over the contextual variables yields estimates of the influence of these variables on offensive and defensive efficiency. The results of the regressions are shown in Tables 6-9. To assess the robustness of our findings related to the post-COVID dummy, Tables 6-9 also present results (i) for a placebo test where estimation is carried out from 2010-11 to 2018-19, and a generalised 2018-19 (placebo) dummy is included for all leagues. This is done to see whether some time-varying factors, unrelated to the COVID-19 outbreak, may drive our results; (ii) for the potential confounding role of the introduction of the VAR, including the series of binary variables described in the previous subsection alongside with the post-COVID dummy[5].

In Table 6 we analyse the impact of our contextual variables on the teams' offensive efficiency in home games. Columns 1 and 2 show that the post-COVID dummy is weakly significant only when the penalties attempted are included in the regression. The placebo test meets with success, as the related dummy is not significant (columns 3-4). The introduction of the VAR dummy, that is also not significant (columns 5-6), does not significantly change our results, as the post-COVID dummy shows a significant impact only in the regression including the penalties attempted. The other contextual variables have the desired sign in all the six columns: *penalties_attempted* has a positive impact on home offensive efficiency, whereas *n_managers* – our proxy for an unlucky season - shows the expected negative sign. Among the country dummies, only France presents a significative negative coefficient in all the regressions. In column 2, we test for a differential league effect of the post-COVID dummy. The null

[5] We stress that this exercise cannot yield a proper counterfactual assessment of the introduction of the VAR, which we are currently carrying out in a companion paper.

hypothesis that ghost matches affect home offensive efficiency uniformly cannot be rejected at a very high significance level.

Table 6 about here

In Table 7 we present the results of the regressions focusing on teams' offensive efficiency in away games. The post-COVID dummy is always strongly significant and positive, therefore playing behind closed doors has increased the offensive efficiency of the visiting teams. The coefficients are higher in regressions including the penalties attempted (columns 2 and 6). The placebo dummy, as well as the VAR dummy, are again irrelevant. The other contextual variables have always the a priori expected sign. The previous results are also confirmed for the league dummies, where only the France coefficient is always significant and negative when penalties are not taken into account, and for the differential league effect, that is not significant.

Table 7 about here

Table 8 shows the results regarding the teams' defensive efficiency in home games. In this case there is not a clear evidence of the impact of ghost games on efficiency, as the post-COVID dummy shows some significance only if penalties are not considered (see columns 1 and 5). Hence, the only relevant changes that have occurred in the post-COVID season are those related to the awarding of penalties. The placebo test was again carried out with success, and the introduction of the VAR does not change the results. The number of penalties converted by the opponents has the expected negative impact on defensive efficiency, and so does the number of managers employed in a season. Among the league dummies, the Ligue 1 coefficient is always significant and negative, whereas the Bundesliga coefficient is significant and negative only when penalties are taken into account. The test for the differential league effect shows some evidence of a differential post-COVID country effect (the full estimates, available upon request, make clear that there is an extra positive impact for Spain).

Table 8 about here

Finally, in Table 9 we can see the results for the teams' defensive efficiency in away games. The post-COVID dummy is always significant and positive, especially in the specifications with penalties (see columns 2 and 6). Therefore, playing behind closed doors has a positive impact on away defensive efficiency. The placebo dummy is once again not significant, just as the VAR dummy. The results related to the other contextual variables are like those obtained when analysing the home defensive efficiency. In particular, the Ligue 1 coefficient is always significant and negative, while the Bundesliga coefficient is significant and negative only alongside with penalties. However, differently from what happens with home defensive efficiency, there is no evidence of a differential league effect: the influence of playing behind closed doors on the defensive efficiency of visiting teams is the same in all countries.

Table 9 about here

In general[6], our evidence shows that in the post-COVID season efficiency increased for away games - offensively, but also defensively. More specifically, the post-COVID dummy is always strongly significant and positive in the regressions focusing on the away offensive efficiency and is significant and positive in the regressions focusing on the away defensive efficiency, especially if penalties are considered. On the other hand, the home offensive efficiency shows a very slight increase, as the post-COVID dummy is weakly significant and positive only with penalties, whereas technical efficiency remains basically the same as before for home defence, as the post-COVID dummy shows some significance only if penalties are *not* considered. These findings are consistent with the recent literature on the impact of ghost games on teams' performance (Endrich and Gesche, 2020; Reade and Singleton, 2020; Reade et al., 2020; Tilp

[6] The regressions presented in Tables 6-9 have also been carried out on the alternative production set with crosses in place of through balls. The results, available upon request, are virtually unchanged. The only difference worth mentioning is that no differential effect exists any longer for the Liga's home defensive efficiency.

and Thaller, 2020; Fischer and Haucap, 2021) and show an erosion of the home advantage also when the analysis is focused on teams' technical efficiency. As evidenced in our literature review, home advantage can be explained by the interaction of three main factors: home crowd support affecting visiting teams' performance and referee's decisions, home teams' familiarity with local playing conditions and visiting teams' travel fatigue. Since the latter two factors are not affected by the ghost games, it is evident that the erosion of the home advantage is determined by the reduced pressure deriving from the lack of home crowd support, which helps visiting teams to be more efficient.

It is important to highlight that the post-COVID dummy affects all leagues equally. Home defensive efficiency is the only case where there is some weak evidence of a differential post-COVID country effect (for Spain), but even this marginal exception vanishes when considering an alternative production set (with crosses replacing through balls). Moreover, Ligue 1 is consistently less efficient than the other countries. A possible explanation relies on the lower quality of French teams, as evidenced by their average roster value over the period under investigation (94.3m euros) in comparison with the other four leagues (254.2m Premier League, 176.4m La Liga, 153.6m Serie A and 146m Bundesliga)[7]. Deloitte's Annual Review of Football Finance (2014-18) also shows how the average team payroll in Ligue 1 (1.11b euros) is the lowest among the top 5 European leagues (2.95b Premier League, 1.63b La Liga, 1.44b Bundesliga and 1.41b Serie A).

The other contextual variables have always the desired sign: *penalties_attempted* has a positive impact on offensive efficiency, *penalties_conceded* a negative impact on defensive efficiency, whereas *n_managers* – our proxy for an unlucky season – always shows the expected negative

[7]These data were collected from the website transfermarkt.it and refer to the roster value at the end of the summer transfer market window.

sign for both types of efficiency. Finally note that the placebo test is always successful and that the VAR dummy never impacts on our findings about the post-COVID season.

5. Conclusions

This study aims to contribute to the recent literature on the effects of COVID on football teams' performance, focusing on the impact of playing behind closed doors – due to the health and safety measures following the COVID-19 pandemic outbreak - on offensive and defensive technical efficiency. Using a long season-level dataset for the top 5 European leagues, a novelty for efficiency studies on football, the analysis compares the ten seasons (2009-10 to 2018-19) played before the pandemic outbreak with the only season (2020-21) entirely played behind closed doors. The methodology applied to calculate the efficiency scores is the conditional order-m, whose application represents a further novel contribution to the literature on football teams' efficiency.

Our results show that in the post-COVID season both offensive and defensive efficiency significantly increased for away games, whereas for home games offensive efficiency shows a very slight increase, and defensive efficiency remains basically unchanged. These results are valid for all the 5 leagues under investigation, as there is no consistent evidence of differential post-COVID country effect. These findings are consistent with the recent literature on the impact of ghost games on teams' performance and show an erosion of the home advantage also when the analysis is focused on teams' efficiency, likely due to the reduced pressure on visiting teams deriving from the lack of home crowd support.

References

Banker R. D. & Morey R. C. (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research*, *34*(4), 513-521.

Barros, C.P., Assaf, A. & Sá-Earp, F. (2010). Brazilian football league technical efficiency: A Simar and Wilson approach. *Journal of Sports Economics*, *11*(6), 641-651.

Barros, C. P. & Douvis, J. (2009). Comparative analysis of football efficiency among two small European countries: Portugal and Greece. *International Journal of Sport Management and Marketing*, *6*(2), 183-199.

Barros, C. P. & Garcia-del-Barrio, P. (2011). Productivity drivers and market dynamics in the Spanish first division football league. *Journal of Productivity Analysis*, *35*(1), 5-13.

Barros, C. P. & Leach, S. (2006). Performance evaluation of the English Premier Foot- ball League with data envelopment analysis. *Applied Economics, 38*(1), 1449–1458.

Barnett, V., & Hilditch, S. (1993). The effect of an artificial pitch surface on home team performance in football (soccer). *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *156*(1), 39-50.

Benz, L. S., & Lopez, M. J. (2021). Estimating the change in soccer's home advantage during the Covid-19 pandemic using bivariate Poisson regression. *AStA Advances in Statistical Analysis*, 1-28.

Boscá, J. E., Liern, V., Martínez, A., & Sala, R. (2009). Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football. *Omega*, *37*(1), 63-78.

Bryson, A., Dolton, P., Reade, J. J., Schreyer, D., & Singleton, C. (2021). Causal effects of an absent crowd on performances and refereeing decisions during COVID-19. *Economics Letters*, *198*, 109664.

Buriamo, B., Paramio, J. L. & Campos, C. (2010). The impact of televised football on stadium attendance in English and Spanish League football. *Soccer & Society*, *11*(4).

Cazals, C., Florens, J. P. & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, *106*(1), 1-25.

Charnes, A., Cooper, W.W. & Rhodes, E. (1981). Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through. *Management Science*, *27*, 668-697.

Clarke, S. R. & Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *44*(4), 509-521.

Coelli, T. J., Rao, D. D. P., O'Donnell, C. J. & Battese, G. E. (2005). *An Introduction to Efficiency and Productivity Analysis* (2nd ed.). New York, USA: Springer.

Correia-Oliveira, C. R., & Andrade-Souza, V. A. (2021). Home advantage in soccer after the break due to COVID-19 pandemic: does crowd support matter? *International Journal of Sport and Exercise Psychology*, 1-12.

Courneya, K. S. & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology*, *14*, 28-39.

Daraio, C. & Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: a probabilistic approach. *Journal of Productivity Analysis*, *24*(1), 93-121.

Daraio, C., Simar, L. & Wilson, P. W. (2018). Central limit theorems for conditional efficiency measures and tests of the 'separability' condition in non-parametric, two-stage models of production. *The Econometrics Journal*, *21*(2), 170-191.

Dávila, C. G. T. & Cebrián, L. I. G. (2012). Eficiencia y resultados deportivos: aplicación a la liga mexicana de fútbol. *Revista Movimiento Humano, 3*, 61-76.

Dawson, P., Dobson, S. & Gerrard, B. (2000). Stochastic frontiers and the temporal structure of managerial efficiency in English soccer. *Journal of Sports Economics*, *1*(4), 341-362.

Deloitte (2014). *Deloitte Annual Review of Football Finance*. Deloitte.

Deloitte (2015). *Deloitte Annual Review of Football Finance*. Deloitte.

Deloitte (2016). *Deloitte Annual Review of Football Finance*. Deloitte.

Deloitte (2017). *Deloitte Annual Review of Football Finance.* Deloitte.

Deloitte (2018). *Deloitte Annual Review of Football Finance*. Deloitte.

De Witte, K. & Kortelainen, M. (2013). What Explains the Performance of Students in a Heterogeneous Environment? Conditional Efficiency Estimation with Continuous and Discrete Environmental Variables. *Applied Economics*, *45*(17), 2401-2412.

Endrich, L. M. & Gesche, T. (2020). Home-bias in referee decisions: Evidence from “Ghost Matches” during the COVID19-Pandemic. *Economics Letters*, *197*, 109621.

Espitia-Escuer, M. & García-Cebrián, L. I. (2004). Measuring the efficiency of Spanish first-division soccer teams. *Journal of Sports Economics*, 5(4), 329–346.

Espitia-Escuer, M. & García-Cebrián, L. I. (2010). Measurement of the efficiency of football teams in the Champions League. *Managerial and Decision Economics*, *31*(1), 373–386.

Ferraresi, M. & Gucciardi, G. (2021). Who chokes on a penalty kick? Social environment and individual performance during COVID-19 times. *Economics Letters*, *203*, 109868.

Fischer, K. & Haucap, J. (2021). Does the crowd support drive the home advantage in professional soccer? Evidence from German Ghost Games during the COVID-19 Pandemic. *Journal of Sports Economics*, 15270025211026552.

García-Sánchez, I. M. (2007). Efficiency and effectiveness of Spanish football teams: A three-stage-DEA approach. *Central European Journal of Operations Research*, *15*(1), 21– 45.

Garicano, L., Palacios-Huerta, I. & Prendergast, C. (2005). Favoritism under social pressure. *The Review of Economics and Statistics*, *87*(2), 208-216.

González-Gómez, F. & Picazo-Tadeo, A. J. (2010). Can we be satisfied with our foot- ball team? Evidence from Spanish professional football. *Journal of Sports Economics*, *11*(4), 418–442.

Goumas, C. (2014). How does crowd support contribute to home advantage in soccer? *Journal of Sports Behavior*, *37*(3), 236.

Haas, D. J. (2003). Productive efficiency of English football teams – A data envelopment analysis approach. *Managerial and Decision Economics*, *24*(1), 403–410.

Hegarty, T. (2021). Information and price efficiency in the absence of home crowd attendance. *Applied Economics Letters*, 1-6.

Jamieson, J. P. (2010). The Home Field Advantage in Athletics: A meta-analysis. *Journal of Applied Social Psychology*, *40*(7), 1819-1848.

Kneip, A., Park, B.U. & Simar, L. (1998). A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econometric Theory*, *14*(6), 783-793.

Kulikova, L. I. & Goshunova, A. V. (2013). Measuring efficiency of professional football club in contemporary research. *World Applied Sciences Journal*, *25*(2), 247-257.

Legaz-Arrese, A., Moliner-Urdiales, D. & Munguia-Izquierdo, D. (2013). Home advantage and sports performance: evidence, causes and psychological implications. *Universitas Psychologica*, *12*(3), 933-943.

Li, Q., & Racine, J. (2007). *Nonparametric econometrics: Theory and practice*. Princeton: Princeton University Press.

Li, Q., & Racine, J. S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business and Economic Statistics*, *26*(4), 423–434.

McCarrick, D., Bilalic, M., Neave, N., & Wolfson, S. (2020). Home Advantage during the COVID-19 Pandemic in European football. *Psychology of Sport and Exercise*, 102013.

Moore, J. C., & Brylinsky, J. (1995). Facility familiarity and the home advantage. *Journal of Sport Behavior*, 18(4), 302-311.

Neale, W. C. (1964). The particular economics of professional sports: A contribution to the theory of the firm in sporting competition and in market competition. *The Quarterly Journal of Sports Economics, 78*(1), February, 1-14.

Nevill, A. M., Newell, S. M. & Gale, S. (1996). Factors associated with home advantage in English and Scottish soccer matches. *Journal of Sports Sciences*, *14*(2), 181-186.

Pollard, R. (1986). Home advantage in soccer: a retrospective analysis. *Journal of Sport Sciences*, *4*, 237-248.

Pollard, R. (2006). Home advantage in soccer: variations in its magnitude and a literature review of the inter-related factors associated with its existence. *Journal of Sport Behavior*, *29*(2), 341-352.

Pollard, R. & Gomez, M. A. (2009). Home advantage in football in South-West Europe: Long-term trends, regional variation, and team difference. *European Journal of Sport Science, 9*(6), 341-352.

Pollard, R. & Pollard, G. (2005). Long-term trends in home advantage in professional team sports in North America and England. *Journal of Sports Sciences*, *23,* 337-350.

Ponzo, M. & Scoppa, V. (2018). Does the home advantage depend on crowd support? Evidence from same-stadium derbies. *Journal of Sports Economics*, *19*(4), 562-582.

Ramchandani, G., & Millar, R. (2021). Investigating the “Twelfth Man” Effect in Five European Domestic Football Leagues: A COVID-19 Induced Natural Experiment. *Journal of Global Sport Management*, 1-15.

Reade, J., Schreyer, D., & Singleton, C. (2020). Eliminating supportive crowds reduces referee bias. *Available at SSRN 3743972*.

Reade, J. J., & Singleton, C. (2020). European football after COVID-19. *A New World Post COVID-19*, 349.

Ribeiro, A. S. & Lima, F. (2012). Portuguese football league efficiency and players' wages. *Applied Economics Letters*, *19*(6), 599-602.

Rossi, G., Goossens, D., Di Tanna, G. L. & Addesa, F. (2019). Football team performance efficiency and effectiveness in a corruptive context: the Calciopoli case. *European Sport Management Quarterly*, *19*(5), 583-604.

Sala-Garrido, R., Carrión, V. L., Esteve, A. M. & Boscá, J. E. (2009). Analysis and evolution of efficiency in the Spanish soccer league. *Journal of Quantitative Analysis in Sports*, *5*(1).

Schwartz, B. & Barksy, S. F. (1977). The home advantage. *Social Forces*, *55*(3), 641-661.

Scoppa, V. (2021). Social pressure in the stadiums: Do agents change behaviour without crowd support? *Journal of Economic Psychology*, 82.

Simar, L. & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, *136*(1), 31-64.

Simar, L. & Wilson, P. W. (2015). Statistical Approaches for Non-parametric Frontier Models: A Guided Tour. *International Statistical Review*, *83*(1), 77–110.

Sloane, P. J. (1971). The economics of professional football: the football club as a utility maximiser. *Scottish Journal of Political Economy*, *18*(2), 121-146.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature, 41*(4), 1137-1187.

Tiedemann, T., Francksen, T. & Latacz-Lohmann, U. (2011). Assessing the performance of German Bundesliga football players: a non-parametric metafrontier approach. *Central European Journal of Operations Research*, *19*(4), 571-587.

Tilp, M., & Thaller, S. (2020). COVID-19 has turned home-advantage into home-disadvantage in the German soccer Bundesliga. *Frontiers in Sports and Active Living*, *2*, 165.

Van Damme, N. & Baert, S. (2019). Home advantage in European international soccer: Which dimension of distance matters? *Economics*, 13.

Villa, G. & Lozano, S. (2016). Assessing the Scoring Efficiency of a Football Match. *European Journal of Operational Research, 255*(1), 559–569.

Wilson, P.W. (2020). *Package 'FEAR'*, August 31, available at https://pww.people.clemson.edu/Software/FEAR/fear.html.

Wunderlich, F., Weigelt, M., Rein, R. & Memmert, D. (2021). How does spectator presence affect football? Home advantages in European top-class football matches played without spectators during the COVID-19 pandemic. *PLOS ONE*.

Zambon Ferraresi, F., Lera López, F. & García Cebrián, L. I. (2017). *Sports results measurement and efficiency in UEFA Champions League* (No. ART-2017-103980).

Table 1. Recent studies on the impact of ghost games on Home Advantage (HA)

Authors	Leagues	Methodology	Metrics	Findings
Benz & Lopez (2021)	13 countries, 17 leagues	Bivariate Poisson regression analysis	Goal difference Yellow card difference	Negative impact of ghost games on HA in some leagues, positive impact in others
Bryson et al. (2021)	17 countries, 23 leagues	Regression analysis	Likelihood of a home win Goal difference Total goals scored Total yellow cards	No impact of ghost games on HA in terms of likelihood of a home win, goal difference and total goals scored Negative impact of ghost games on HA in terms of yellow cards
Correia-Oliveira & Andrade-Souza (2021)	First and second division of English, German, Italian and Spanish leagues	Descriptive and correlation analysis	Ratio between the number of points won at home and the total number of points won at home and away	Negative impact of ghost games on HA
Endrich and Gesche (2020)	First and second divisions of the German league	Regression analysis	Total fouls Total yellow cards	Negative impact of ghost games on HA
Ferraresi & Gucciardi (2021)	First division of English, French, German, Italian and Spanish leagues	Regression analysis	Probability of missing a penalty	Negative impact of ghost games on HA
Fischer & Haucap (2021)	First, second and third division of the German league	Regression analysis	Home wins Difference in points earned by home and away team	Negative impact of ghost games on HA
McCarrick et al. (2020)	11 countries, 15 leagues	Multilevel regression analysis	Total points Total goals scored Total fouls Total yellow cards	Negative impact of ghost games on HA in relation to all the indicators, but team dominance diluted this impact

Ramchandani & Miller (2021)	First division of English, German, Italian, Portuguese and Spanish leagues	Descriptive analysis	Ratio between the number of points won at home and the total number of points won at home and away	Negative impact of ghost games on HA only in German and Italian leagues
Reade et al. (2020)	UEFA Champions League and Europa League French Ligue 1 First, second and third division of the Italian league Coppa Italia	Regression analysis	Goal difference Total yellow cards Yellow card difference Yellow cards per foul	No significant impact of ghost games on HA in terms of goal difference Negative impact of ghost games on HA in terms of yellow cards
Reade & Singleton (2020)	German first division	Descriptive analysis	Percentage of home and away wins	Negative impact of ghost games on HA
Scoppa (2021)	First and second divisions of English, German, Italian, Portuguese and Spanish leagues	Regression analysis	Total points Total goals scored Total shots Total shots on target Total corner kicks Total fouls Total yellow and red cards Total penalties	Negative impact of ghost games on HA in relation to all the indicators
Tilp & Thaller (2020)	German first division	Descriptive analysis	Total fouls Total yellow and red cards Total penalties	Negative impact of ghost games on HA in relation to all the indicators

Table 2. Offensive inputs and outputs and contextual variables

Variable	Type	Description
Points	Offensive/defensive output	number of points per game
goals_scored	Offensive output	number of goals scored per game
inv_goals_conceded	Defensive output	inverse of goals conceded per game
shots_made, possession, dribbles, through_balls	Offensive input	shot attempts per game
Possession	Offensive/defensive input	possession percentage per game
Dribbles	Offensive input	successful dribbles per game
through_balls	Offensive input	through balls per game
Crosses	Offensive input	crosses per game
inv_shots_conceded	Defensive input	inverse of opponents' shot attempts per game
Tackles	Defensive input	tackles per game
inv_through_balls_conceded	Defensive input	inverse of opponents' through balls per game
inv_crosses_conceded	Defensive input	inverse of opponents' crosses per game
penalties_attempted	Contextual variable	number of penalties attempted per game
penalties_conceded	Contextual variable	number of opponents' penalties converted per game
n_managers	Contextual variable	number of managers employed in a season
post_COVID	Contextual variable	2020-21 season binary variable
Placebo	Contextual variable	placebo binary variable (see text for details)
V AR	Contextual variable	VAR introduction binary variables (see text for details)
Ligue 1, Bundesliga, Serie A, La Liga	Contextual variables	league binary variables

Table 3. Main descriptive statistics, No. of observations=1078

	Home						Away					
	Mean		Standard Deviation		Median		Mean		Standard Deviation		Median	
Variables	**Pre-COVID**	**Post-COVID**	**Pre-COVID**	**Post-COVID**	**Pre-COVID**	**Post-COVID**	**Pre-COVID**	**Post-COVID**	**Pre-COVID**	**Post-COVID**	**Pre-COVID**	**Post-COVID**
points	1.64	1.45	0.49	0.50	1.58	1.39	1.11	1.29	0.47	0.52	1.00	1.21
goals_scored	1.55	1.48	0.54	0.56	1.42	1.47	1.16	1.32	0.43	0.42	1.05	1.21
shots_made	14.29	12.63	2.49	2.38	13.90	12.20	11.45	11.08	2.07	2.17	11.20	10.55
possession	50.88	50.78	4.42	4.97	50.20	50.50	49.11	49.22	4.72	5.00	48.50	49.20
dribbles	9.03	9.56	2.73	1.90	8.70	9.55	8.36	9.21	2.63	1.88	8.05	9.25
through_balls	2.16	0.56	1.86	0.52	2.00	1.00	1.87	0.52	1.62	0.56	1.00	0.20
crosses	22.70	19.06	4.40	2.96	23.00	19.00	17.90	17.10	3.59	2.72	18.00	17.00
goals_conceded	1.16	1.32	0.35	0.34	1.16	1.32	1.55	1.48	0.42	0.44	1.53	1.47
shots_conceded	11.45	11.08	2.03	2.10	11.50	11.15	14.25	12.63	2.40	2.25	14.20	12.60
tackles	19.42	15.08	2.62	1.85	19.35	14.85	19.69	15.38	2.57	1.91	19.70	15.30
through_balls_conceded	1.87	0.69	1.29	0.39	1.30	0.49	2.14	0.67	1.46	0.40	1.47	0.49
crosses_conceded	17.89	17.13	3.48	2.93	18.00	17.00	22.52	19.09	4.33	3.11	23.00	19.00
penalties_attempted	0.17	0.20	0.11	0.13	0.16	0.16	0.11	0.18	0.08	0.11	0.11	0.16
penalties_conceded	0.04	0.07	0.03	0.05	0.03	0.08	0.07	0.08	0.04	0.05	0.06	0.08
n_managers	1.48	1.44	0.74	0.75	1.00	1.00	1.48	1.44	0.74	0.75	1.00	1.00

Table 4. Efficiency scores: Offense

Outputs: goals_scored, points

Inputs: possession, shots_made, dribbles, through_balls

No. of observations=1078

Order-m	**Cond. order-m with league dummies, n_managers, post-COVID dummy**	**Cond. order-m with league dummies, n_managers, post-COVID dummy, penalties_attempted**
	Home	
Min. :0.8150 1st Qu.:1.0460 Median :1.2277 Mean :1.3074 3rd Qu.:1.4620 Max. :3.6363	Min. :0.8911 1st Qu.:1.0000 Median :1.0995 Mean :1.1969 3rd Qu.:1.3104 Max. :3.4771	Min. :0.9965 1st Qu.:1.0000 Median :1.0000 Mean :1.0713 3rd Qu.:1.0666 Max. :2.4091
	Away	
Min. :0.8307 1st Qu.:1.0269 Median :1.2509 Mean :1.3428 3rd Qu.:1.5262 Max. :3.3000	Min. :0.9164 1st Qu.:1.0000 Median :1.0964 Mean :1.2199 3rd Qu.:1.3511 Max. :2.6760	Min. :0.9943 1st Qu.:1.0000 Median :1.0000 Mean :1.0878 3rd Qu.:1.1059 Max. :2.3684

Table 5. Efficiency scores: Defence

Outputs: inv_goals_conceded, points

Inputs: possession, inv_shots_conceded, tackles, inv_through_balls_conceded

No. of observations=1078

Order-m	**Cond. order-m with league dummies, n_managers, post-COVID dummy**	**Cond. order-m with league dummies, n_managers, post-COVID dummy, penalties_conceded**
Home		
Min. :0.7755 1st Qu.:1.0000 Median :1.2047 Mean :1.2761 3rd Qu.:1.4354 Max. :2.5319	Min. :0.8833 1st Qu.:1.0000 Median :1.0618 Mean :1.1726 3rd Qu.:1.2771 Max. :2.3051	Min. :0.9937 1st Qu.:1.0000 Median :1.0000 Mean :1.0824 3rd Qu.:1.0905 Max. :2.1250
Away		
Min. :0.8187 1st Qu.:1.0000 Median :1.2175 Mean :1.3027 3rd Qu.:1.4832 Max. :3.8025	Min. :0.9185 1st Qu.:1.0000 Median :1.0476 Mean :1.1767 3rd Qu.:1.2727 Max. :2.9954	Min. :0.9988 1st Qu.:1.0000 Median :1.0000 Mean :1.0686 3rd Qu.:1.0405 Max. :2.6027

Table 6. Regressions from conditional order-m: home offensive efficiency

No. of observations=1078

Dependent variable in columns (1), (3), (5): ratio between home offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, and home offensive efficiency from order-m.

Dependent variable in columns (2), (4), (6): ratio between home offensive efficiency from conditional order-m with league dummies, n_managers, postCOVID dummy, penalties, and home offensive efficiency from order-m.

	(1)	(2)	(3)	(4)	(5)	(6)
Ligue 1	-0.094***	-0.046***	-0.086***	-0.033**	-0.095***	-0.043***
	(0.009)	(0.013)	(0.010)	(0.014)	(0.009)	(0.010)
Bundesliga	0.012*	0.002	0.013	0.002	0.011	0.003
	(0.007)	(0.013)	(0.008)	(0.014)	(0.007)	(0.012)
Serie A	0.006	-0.015	0.007	-0.013	0.005	-0.013
	(0.007)	(0.013)	(0.008)	(0.012)	(0.008)	(0.012)
La Liga	-0.004	-0.002	-0.008	-0.000	-0.004	0.000
	(0.007)	(0.014)	(0.008)	(0.012)	(0.007)	(0.010)
n_managers	-0.112***	-0.109***	-0.111***	-0.106***	-0.112***	-0.108***
	(0.005)	(0.009)	(0.005)	(0.009)	(0.005)	(0.009)
post_COVID	0.008	0.029*			0.003	0.044***
	(0.010)	(0.016)			(0.012)	(0.016)
penalties_attempted		0.315***		0.331***		0.317***
		(0.051)		(0.035)		(0.040)
Placebo			0.010	0.001		
			(0.009)	(0.013)		
VAR					0.005	-0.017
					(0.009)	(0.015)
F-test for differential league effect (p-value)		0.661				

Notes. Bootstrapped (2000 replications) standard errors in parentheses. * indicates a coefficient p-value < 0.10, ** a p-value < 0.05, *** a p-value < 0.01. The dummy for the Premier League was omitted as we consider it as the default league. The null hypothesis for the F-test to determine whether there is a differential post-COVID league effect is: post_Ligue 1 = post_Bundesliga = post_Serie A = post_La Liga = 0, where post-Ligue 1, etc., are interactive terms between the league and the post-COVID dummies.

Table 7. Regressions from conditional order-m: away offensive efficiency

No. of observations=1078

Dependent variable in columns (1), (3), (5): ratio between away offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, and home offensive efficiency from order-m.

Dependent variable in columns (2), (4), (6): ratio between away offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, penalties, and home offensive efficiency from order-m.

	(1)	(2)	(3)	(4)	(5)	(6)
Ligue 1	-0.054***	-0.024	-0.058***	-0.017	-0.055***	-0.027*
	(0.011)	(0.015)	(0.012)	(0.016)	(0.011)	(0.014)
Bundesliga	0.008	0.013	0.014	0.021*	0.007	0.008
	(0.009)	(0.011)	(0.010)	(0.013)	(0.009)	(0.015)
Serie A	0.001	-0.012	0.002	-0.003	-0.001	-0.016
	(0.008)	(0.012)	(0.010)	(0.012)	(0.009)	(0.013)
La Liga	-0.001	0.005	0.002	0.014	-0.002	-0.004
	(0.009)	(0.011)	(0.010)	(0.015)	(0.009)	(0.012)
n_managers	-0.093***	-0.116***	-0.095***	-0.112***	-0.093***	-0.115***
	(0.006)	(0.008)	(0.006)	(0.011)	(0.006)	(0.008)
post_COVID	0.035***	0.066***			0.028**	0.046***
	(0.010)	(0.011)			(0.014)	(0.016)
penalties_attempted		0.208***		0.151**		0.197***
		(0.062)		(0.060)		(0.057)
placebo			0.016	0.016		
			(0.011)	(0.016)		
VAR					0.008	0.021*
					(0.011)	(0.012)
F-test for differential league effect (p-value)		0.680				

Notes. Bootstrapped (2000 replications) standard errors in parentheses. * indicates a coefficient p-value < 0.10, ** a p-value < 0.05, *** a p-value < 0.01. The dummy for the Premier League was omitted as we consider it as the default league. The null hypothesis for the F-test to determine whether there is a differential post-COVID league effect is: post_Ligue 1 = post_Bundesliga = post_Serie A = post_La Liga = 0, where post-Ligue 1, etc., are interactive terms between the league and the post-COVID dummies.

Table 8. Regressions from conditional order-m: home defensive efficiency

No. of observations=1078

Dependent variable in columns (1), (3), (5): ratio between away offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, and home defensive efficiency from order-m.

Dependent variable in columns (2), (4), (6): ratio between away offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, penalties, and home defensive efficiency from order-m.

	(1)	(2)	(3)	(4)	(5)	(6)
Ligue 1	-0.068***	-0.045***	-0.059***	-0.047***	-0.068***	-0.051***
	(0.009)	(0.012)	(0.010)	(0.013)	(0.009)	(0.012)
Bundesliga	0.003	-0.060***	0.001	-0.061***	0.004	-0.063***
	(0.008)	(0.013)	(0.009)	(0.015)	(0.008)	(0.014)
Serie A	0.008	-0.011	0.006	0.010	0.009	0.009
	(0.008)	(0.010)	(0.009)	(0.011)	(0.008)	(0.010)
La Liga	0.008	-0.003	0.004	-0.009	0.008	-0.005
	(0.008)	(0.011)	(0.009)	(0.012)	(0.008)	(0.011)
n_managers	-0.095***	-0.096***	-0.093***	-0.092***	-0.095***	-0.095***
	(0.005)	(0.007)	(0.006)	(0.008)	(0.005)	(0.009)
post_COVID	-0.025***	0.001			-0.022*	-0.011
	(0.009)	(0.015)			(0.012)	(0.018)
penalties_conceded		-1.278***		-1.269***		-1.286***
		(0.100)		(0.134)		(0.115)
placebo			-0.002	0.008		
			(0.010)	(0.012)		
VAR					-0.003	0.012
					(0.009)	(0.014)
F-test for differential league effect (p-value)		0.063				

Notes. Bootstrapped (2000 replications) standard errors in parentheses. * indicates a coefficient p-value < 0.10, ** a p-value < 0.05, *** a p-value < 0.01. The dummy for the Premier League was omitted as we consider it as the default league. The null hypothesis for the F-test to determine whether there is a differential post-COVID league effect is: post_Ligue 1 = post_Bundesliga = post_Serie A = post_La Liga = 0, where post-Ligue 1, etc., are interactive terms between the league and the post-COVID dummies.

Table 9. Regressions from conditional order-m: away defensive efficiency

No. of observations=1078

Dependent variable in columns (1), (3), (5): ratio between away offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, and away defensive efficiency from order-m.

Dependent variable in columns (2), (4), (6): ratio between away offensive efficiency from conditional order-m with league dummies, n_managers, post-COVID dummy, penalties, and away defensive efficiency from order-m.

	(1)	(2)	(3)	(4)	(5)	(6)
Ligue 1	-0.072***	-0.057***	-0.074***	-0.053***	-0.071***	-0.056***
	(0.011)	(0.012)	(0.013)	(0.015)	(0.011)	(0.015)
Bundesliga	-0.011	-0.130***	-0.008	-0.134***	-0.010	-0.131***
	(0.010)	(0.016)	(0.011)	(0.013)	(0.010)	(0.016)
Serie A	0.011	0.001	0.012	0.009	0.012	0.003
	(0.008)	(0.013)	(0.010)	(0.012)	(0.009)	(0.012)
La Liga	-0.008	-0.013	-0.010	-0.012	-0.007	-0.013
	(0.010)	(0.014)	(0.011)	(0.014)	(0.010)	(0.013)
n_managers	-0.093***	-0.088***	-0.091***	-0.082***	-0.093***	-0.087***
	(0.006)	(0.010)	(0.006)	(0.009)	(0.006)	(0.009)
post_COVID	0.019*	0.045***			0.024*	0.047***
	(0.010)	(0.013)			(0.013)	(0.018)
penalties_conceded		-0.712***		-0.748***		-0.720***
		(0.097)		(0.134)		(0.107)
placebo			-0.004	-0.020		
			(0.012)	(0.018)		
VAR					-0.007	-0.003
					(0.009)	(0.012)
F-test for differential league effect (p-value)		0.697				

Appendix

<u>Efficiency scores: Offense</u>

Outputs: goals_scored, points.

Inputs: possession, shots_made, dribbles, crosses

No. of observations=1078

Order-m	**Cond. order-m with league dummies, n_managers, post-COVID dummy**	**Cond. order-m with league dummies, n_managers, post-COVID dummy, penalties_attempted**
Home		
Min. :0.8087 1st Qu.:1.0421 Median :1.2466 Mean :1.3273 3rd Qu.:1.4933 Max. :4.0469	Min. :0.8908 1st Qu.:1.0000 Median :1.0989 Mean :1.2107 3rd Qu.:1.3297 Max. :3.9594	Min. :0.9985 1st Qu.:1.0000 Median :1.0000 Mean :1.0789 3rd Qu.:1.0726 Max. :3.3334
Away		
Min. :0.8281 1st Qu.:1.0725 Median :1.3225 Mean :1.4369 3rd Qu.:1.6676 Max. :4.0186	Min. :0.9054 1st Qu.:1.0000 Median :1.1410 Mean :1.2824 3rd Qu.:1.4437 Max. :3.0235	Min. :0.9925 1st Qu.:1.0000 Median :1.0000 Mean :1.1099 3rd Qu.:1.1304 Max. :3.0663

Efficiency scores: Defence

Outputs: inv_goals_conceded, points.

Inputs: possession, inv_shots_conceded, tackles, inv_crosses_conceded

No. of observations=1078

Order-m	**Cond. order-m with league dummies, n_managers, post-COVID dummy**	**Cond. order-m with league dummies, n_managers, post-COVID dummy, penalties_conceded**
Home		
Min. :0.7887 1st Qu.:1.0528 Median :1.2376 Mean :1.3049 3rd Qu.:1.4707 Max. :2.8529	Min. :0.8868 1st Qu.:1.0000 Median :1.1156 Mean :1.2063 3rd Qu.:1.3306 Max. :2.3817	Min. :0.9923 1st Qu.:1.0000 Median :1.0006 Mean :1.1084 3rd Qu.:1.1569 Max. :2.1250
Away		
Min. :0.7908 1st Qu.:1.0899 Median :1.3308 Mean :1.3984 3rd Qu.:1.6081 Max. :3.4661	Min. :0.8882 1st Qu.:1.0000 Median :1.1443 Mean :1.2486 3rd Qu.:1.4098 Max. :2.9710	Min. :0.9898 1st Qu.:1.0000 Median :1.0000 Mean :1.1016 3rd Qu.:1.1322 Max. :2.3714