



## **Python for Economics and Social Sciences**

**To what extent does the presence of supporters  
conditionate the winning probabilities of the home-  
team in soccer?**

**Jean Fontalirand, Colombar Bilbault, Javier Oest Fernández, Simon  
Rühle**

December 10, 2021

Lecturer: Moritz Pfeifer

## Introduction

On June 24th 2021 the **UEFA** (Union of European Football Associations), announced [its decision to abolish the away goals rule in all UEFA club competitions](#) as of the 21/22 season. This rule was applied to determine the winner of a two-legged knockout tie in cases where the two teams had scored the same number of goals on aggregate over the two matches. In such cases, the team which had scored the higher number of goals away from home was considered the winner of the tie and qualified for the next round of the competition.

It was first applied in the UEFA Cup Winners' Cup when Budapest Honvéd beat Dukla Prague in the second round in 1965–66 and has accompanied us personally throughout our childhood and youth and either helped our favourite teams to qualify for the next round or led to their elimination from the competition. It was founded on the presumption that playing at home would be beneficial in comparison to playing away. The inconvenience of travelling for hours, of having to stay several nights abroad, of not being used to the field's characteristics, as well as the presence of the opposite spectators played hardly against the away teams. That is why the proportion of home/away wins was of 61%/19%, which, according to the UEFA has now been reduced to 47%/30% (it does not add up to 100% because we also must consider the draws) In order to justify its decision and to provide an explanation for the gap's reduction, the UEFA made the following statement:

'Many different factors may be considered as having an impact on this decline in home advantage. Better pitch quality and standardised pitch sizes, improved stadium infrastructure, higher security conditions, enhanced care of refereeing (and more recently the introduction of technological support such as GLT and VAR), wider and more sophisticated TV coverage of matches, more comfortable travel conditions, a compressed calendar dictating squad turnover, and changes in competition formats are all elements which have affected the way football is played and blurred the lines between playing at home and away.'

However, there is a factor that has not changed at all during the past 60 years: **supporters**. Although pitch and travel conditions may have improved, away teams have always had to face the thunderous roar of the home supporters who overwhelmingly exceed the number of away fans who are able to travel to support their team and are allowed into the stadiums.

As all other factors can now be considered as equal between home or away, we wondered what impact supporters would have on the results of their teams.

Thanks to the *COVID-19-crisis*, a door opened for us. With soccer competitions being held behind closed doors during 6 to 12 months between 2020 and 2021, we disposed of a large sample of matches without spectators to compare with the ones in which home teams had been able to feel the warmth and the support of their fans.

Several studies have looked at this phenomenon and have also taken advantage of the pandemic period to verify the impact of spectators on the advantage held by the home team. Two of them have particularly caught our attention because of their quite different results.

On the one hand, Carlos Cueva<sup>1</sup>, a Spanish economist, analysed 233,666 games played since 1993-1994 in 41 leagues. Of these games, 2,749 took place behind closed doors during the pandemic. He found a clear decrease in the share of home wins in the total number of games in the period during which games were played without spectators (from 45% to 41%). The main

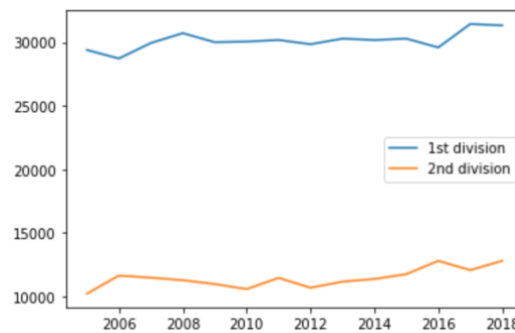
---

<sup>1</sup> Cueva, Carlos, 2020. "Animal Spirits in the Beautiful Game. Testing social pressure in professional football during the COVID-19 lockdown," OSF Preprints hczkj, Center for Open Science.  
<https://ideas.repec.org/p/osf/osfxxx/hczkj.html>

explanation he gives is the pressure that the crowd would have on the referee's behaviour, who would be more likely to favour the home team in the presence of the public. On the other hand, German researchers from the University of Cologne<sup>2</sup> found a different result. By analysing about 37,000 matches, including a thousand played behind closed doors, they found no real difference between the percentage of home wins before and after the pandemic. The percentage of home wins in total games played only drops from 45% to 43%. While they did not find that the outcome of the games was fundamentally different, they did show that the absence of spectators changed other aspects of the game: for example, the fact that the referee gave fewer yellow and red cards to the visiting team in the absence of spectators.

This difference in results made us wonder: how can we really determine the impact of the spectators on the advantage given to the home team?

Although we understand the result may change depending on the sample of data we select, we wanted to conduct our own study. We also wanted to ask ourselves another question: is there a difference between the first and second division? Indeed, as we calculated, the average attendance is much lower in the minor divisions, concretely 38%.



However, stadiums in the second divisions are on average smaller, so less fans bring a more intense atmosphere. It would have been more valuable to analyse the percentage of attendance in the stadium, but we could not find any data about it. Even though a more intense atmosphere should be more important than the absolute number of fans, we made the hypothesis that spectators in the second league, considering their lower attendance, have less impact on the home team's performance.

Thus, **to what extent does the presence of the supporters conditionate the winning probabilities of the home-team in soccer?** In other words, can we observe a difference between the percentage of victories of the home teams during the period with and without spectators? **And if there is a difference, is there a gap between the first and second divisions?**

## The data

To answer this question and to find out if having spectators in the stadium or the fact playing at home or away is relatively more important to win a match depending on the division, we analysed datasets of the following leagues from the season 2005-2006 to the season 2020-2021.

<sup>2</sup> Wunderlich F, Weigelt M, Rein R, Memmert D (2021) How does spectator presence affect football? Home advantage remains in European top-class football matches played without spectators during the COVID-19 pandemic. PLoS ONE 16(3): e0248590. <https://doi.org/10.1371/journal.pone.0248590>

- France: Ligue 1, Ligue 2
- Italy: Serie A, Serie B
- Spain: LaLiga, Segunda División
- England: Premier League, Championship
- Germany: Bundesliga, Bundesliga 2

To calculate the attendance average and the difference between both divisions, we used the data available at <http://www.european-football-statistics.co.uk/attn.htm>.

The results of the matches, as well as the scored and conceded goals came from <https://www.football-data.co.uk/>.

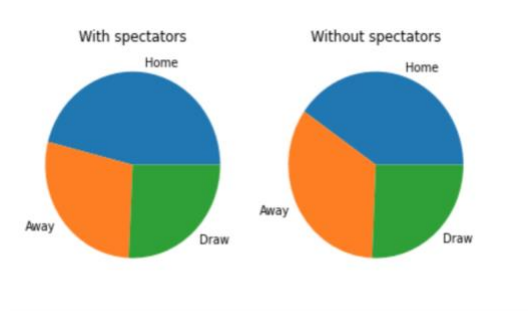
The data we collected amounted to 28,735 1st-league-match results (with home and away goals) and 33,656 results from 2nd-League matches, that is, a **total of 62,391 fixtures**. Among them, there were 1,343 1st-league-matches without spectators and 1,310 2nd-league-ghost matches, for a total of **2,653 games played behind closed doors**.

We downloaded all the files into our computers and classed them into different folders. Hereafter, we had to determine the time during which no spectators were present in the stadiums due to the national measures against Covid-19. While the competitions were suspended across all leagues around the 11th March 2020 when the OMS officially declared [COVID-19 as a pandemic](#), matches were gradually and irregularly resumed across countries from the following season onwards. As there was no indicator in our datasets referring to the spectators' attendance, we manually marked the ghost games by assigning them an additional variable, 'no'. Having done this, we could begin with the coding to analyse the downloaded data.

## Results

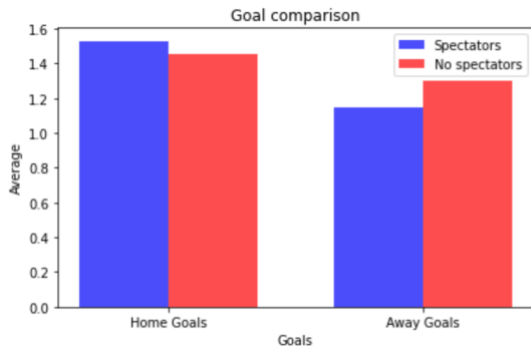
The first observation that can be made from the analysis of our data is that home teams won less during the period without spectators than during the period with spectators. To compare these percentages, we created two different dataframes: one gathering the games played with spectators and the other gathering the games played without spectators. After that, we calculated the percentage of wins of the home team in each of these dataframes. We found that in the period before the pandemic, the home teams won 46.02% of the matches. During the pandemic, without spectators, they only won 40.3% of the games, which is 5.7 percentage points less than during the period with spectators.

Looking at how these percentage points were redistributed proves to be very interesting. The number of draws did practically not change: 25.71% during the non-covid period, against 25.68% of the matches in the period when the stadiums were closed to the public. The entire fall in home-wins converted into away-wins. That means that home teams lost 5.7% more without the support of their fans.



This first observation is already a part of the answer to our question: Yes, home teams are, according to our statistics, more advantaged when there are spectators in the stands. Our database also allowed us to compare the number of goals conceded and scored by each team in each game.

An average game would have a score of 1.72-1.20 with spectators in the stadiums. If we compare it with the average score of ghost matches, we can highlight a remarkable rapprochement: It passes to 1.6-1.39, as can be seen in the graph below:



	No spectators	Spectators
Home Goals	1,6	1,72
Away Goals	1,39	1,2

This is along the same lines as the results we mentioned before concerning the redistribution of the winning probability. When the difference between the goals scored between the home and the away team is higher, like it is the case in times with spectators, the home team is more likely to win.

As a partial, general conclusion for this part we can say that the lack of the fans in the stands notably influences the first division teams: Not only less goals are scored but also more goals are conceded, therefore reducing the chance of winning and increasing that of losing.

What can we notice when we conduct the same analysis on the second division?

First, we realise that in the period with spectators the home advantage is less important in the second division. Indeed, while the home teams win nearly 46.02% of the matches in the first division, the second division teams playing at home do alike only 44.3% of the time. This difference, even if it is not very significant (only 1.72%) could be a first hint to confirm our starting assumption: the more spectators there are, the greater the home advantage.

For the second division, the share of home wins decreased from 44.3% to 40.68%. We can recall that for the first division, this share went from 46.02% to 40.28%. That is, the effect of losing the spectators for the home team was less striking in the second leagues than in the first leagues.

Furthermore, one result particularly struck us. Between the period with and without spectators, the share of victories of teams playing away does not increase much in the second division (from 26.2% to 27% of the total number of games) but increases much more in the first division (from 28.2% to 34.02%). Hence, we can conclude that the greater fall in home wins in the first division is rather compensated by loses, while the same, minor decrease in the second leagues is translated into more draws.

	1st leagues	2nd leagues
$\Delta$ Home wins	-5.741438	-3.636850
$\Delta$ Draws	-0.030432	2.815025
$\Delta$ Away wins	5.771870	0.821825

Again, the fact of winning less (than in the second league) and losing more (while in the second league they draw more) in the absence of fans is another sign that confirms our first hypothesis, that fan presence would be more important for teams in the first division.

## Conclusion

*"Der Fan ist der 12. Mann... but mostly in the top-leagues"*, as it is usually said in German. Indeed, our studies allow us to affirm that the public has a considerable influence on the result of first league matches, since with spectators the team playing at home has a 46.02% chance of winning the match, whereas in COVID period, i.e., without spectators, the probability of the home-team winning was 40.3%. This difference is reflected in the number of goals scored and received against the away team. On average 1.72 goals were scored against the away team and 1.2 goals were received by the home team when there were spectators and 1.59 and 1.39 respectively when there were no spectators. This is in line with what the Spaniard Carlos Cueva concluded in the study he conducted in September 2020, with a different data set and with an additional qualitative approach using a combination of data analysis and psychology: "[M]y findings show a sizable and significant reduction in the home advantage during the lockdown".

What is completely new in our survey is the comparative analysis on the impact of the spectators between first and second leagues. Despite our starting assumption that there would be a big difference between the two divisions due to the different match attendance and the fact that referees in the second league face less pressure, we noticed that the gap, even though it existed, was not so large. However, in the second leagues the fall in wins was compensated by more draws, whereas in the first leagues the ties remained identical and it were the losses that rose.

It should be noted that our study focuses on the impact that the public has and not the way in which it is produced, since we conducted a quantitative and not a qualitative study. Nevertheless, we will briefly review the possible reasons for this impact. Our findings can be due to the encouragement of the team, and the overall ambiance in big, filled stadiums with three times more spectators, but also to the pressure that is put on referees, as was indicated by Cueva (2020) with respect to the first leagues. We could imagine that in a second league game, which is less transcendental and therefore less mediatized, referees would whistle more neutrally. Furthermore, it is possible that the general belief that home teams have an advantage leads to a placebo effect and becomes a self-fulfilling prophecy.

Considering that it was possible to hear the conversations on the field and between the referee and the players during the time without fans in the stadium, it would be interesting to conduct further research and analyse the speeches on the field with Python. By doing so, the hypothesis that higher social pressure is exerted on referees in times with spectators than in periods without spectators could be proven or refuted. We would be curious to have access to this data, but unfortunately it does not exist to date. A possible way of collecting it would be by attaching microphones to the referees during the matches.