How are goals born? External and internal factors in European soccer

¿Cómo nacen los goles? Factores externos e internos en el fútbol europeo

José M. Lavín[1]

Álvaro Jiménez-Sánchez[2]

Aser Falagán-Madrazo[3]

[1] Universidad Internacional de Valencia, Valencia, Spain

[2] Universidad de Valladolid, Valladolid, Spain

[3] Universidad del País Vasco, Lejona, Spain

Correspondence:
Álvaro Jiménez Sánchez
alvarojs@uva.es

José María Lavín de la Cavada
josemaria.lavin@universidadviu.es

Short title:
How are goals born?

How to cite this article:
Lavín, J.M., Jiménez-Sánchez, Á., & Falagán-Madrazo, A. (2023). How are goals born? External and internal factors in European soccer. *Cultura, Ciencia y Deporte, 18*(56), 95-117. https://doi.org/10.12800/ccd.v18i56.1926

Received: 20 May 2022 / ***Accepted:*** 16 January 2023

Abstract

This research studied by means of linear regression models which factors that most influence the achievement of points and goals in eight European leagues in the 2019-2020 season. In order to analyse the points obtained, external variables such as budget, stadium attendance, squad experience, fair play, or whether the players are foreign, national, home-grown or belong to an international team, were taken into account. For the goals, internal or match-specific variables such as shots, scoring chances, corner kicks, number of cards, faults, ball possession, dribbles, tackles, presence or absence of crowd and the pitch factor, were used. The results revealed the importance of budget and fair play, as well as scoring chances and shots. The presence of the public and the pitch factor had hardly any influence. The external factors show regression models with an adjusted R2 between 0.508 and 0.916, while in the case of internal factors range between 0.289 and 0.464. The data may be of relevance for improving transfer investment strategies, as well as for increasing the performance of training sessions, tactics and line-ups.

Keywords: European football, budget, fair play, public, points.

Resumen

Esta investigación estudia mediante modelos de regresión lineal cuáles son los factores que más influyen en la consecución de puntos y goles en ocho ligas europeas de la temporada 2019-2020. Para analizar los puntos obtenidos se tomaron en cuenta variables externas como el presupuesto, asistencia al estadio, experiencia de la plantilla, deportividad, o si los jugadores son extranjeros, nacionales, canteranos o pertenecientes a una selección internacional. Para los goles se recurrió a variables internas o propias de cada partido como los tiros, ocasiones de gol, saques de esquina, número de tarjetas, faltas, posesión del balón, regates, entradas, presencia o ausencia de público y el factor campo. Los resultados revelaron la importancia del presupuesto y de la deportividad, así como de las ocasiones de gol y los tiros. Apenas influye la presencia de público y el factor campo. Los factores externos muestran modelos de regresión con un R2 ajustado entre 0,508 y 0,916, mientras que los internos van entre 0,289 y 0,464. Los datos pueden ser de relevancia para mejorar las estrategias de inversión en fichajes, así como para incrementar el rendimiento de los entrenamientos, las tácticas y las alineaciones.

Palabras clave: fútbol europeo, presupuesto, deportividad, público, puntos.

Introduction

Goals are the most common measure to assess effectiveness in competitions (Hughes & Franks, 2005), but they alone do not measure the performance of a team (Lago-Ballesteros et al., 2012). This is because scores in football are lower than in most other sports, so goals are rare event within the match (Peña & Navarro, 2016). However, in the eye of the fan, they become the almost unique way of understanding the development of the game. Sometimes, the fan forgets external aspects such as the importance of the team's budget or internal characteristics of the game itself, such as the number of goal shooting, ball possession and or faults.

The different relevance of these factors will depend, in turn, on the tradition of the teams, their countries and their historical styles of play (Sarmento et al., 2014) and even other political or social circumstances (Hamil et al., 2010).

Since the end of the 1990s, football has undergone transformations that have permanently altered club competitions, both national and international. Likewise, the championships for national teams have also suffered consequences due to these changes, but perhaps they are more hidden.

The first of these was the so-called Bosman Rulling in 1995, which allowed footballers from any country in the European Union to play as nationals in any another championship of a country in the Union. This fact led to a major restructuring in most of the European professional teams, signing foreign but community players and in some cases relegating national players or the country itself. (Poli et al., 2016). Though, UEFA has subsequently tried to balance this proportion with new regulations (2006-2007 with the quotas of "local" players (Bullough & Jordan, 2017)).

Parallel to this change, most football clubs have ceased to be proper clubs, to become public limited sports companies, with the economic needs that this poses. While a football club also needs money, a public limited company forces you to seek multiple incomes and introduce continuous sources of added value (Andrikopoulos & Kaimenakis, 2006). To this must be added the social prestige and recognition offered by the football showcase.

This seeking of incomes also led to the entry of millionaires. Some of them are from Eastern Europe as in the case of Chelsea, from the United States at Manchester United, the Far East as Inter Milan or from Arab Countries as at Manchester City and Paris Saint Germain, just to name the most relevant. All of them seek this prestige to improve their image (García-Vega, 19 January, 2020), but also an economic return (Solberg & Mehus, 2014).

The third factor of change has also been economic (López-Busto et al., 2016). The arise of satellite television has made football the most global sport and all the national championships are broadcast by a multitude of operators, moving enormous sums of money of which a percentage goes to clubs.

Another source of financing is the influx of public to the stadium. The big football coliseums have proven to be an almost untapped source of income, but, in addition, it is of great importance when it comes to determining the aggressiveness of local teams according to different studies. (Carolina-Paludo et al., 2020; Fothergill et al., 2017; Neave & Wolfson, 2003). The greater the influx of public, the local teams feel more supported and achieve better results due to the pressure exerted by the identification of the followers with the team (Katz et al, 2020). However, it should be mentioned that the latest studies of home pitch advantage raise some doubts about this bystander (Jiménez-Sánchez & Lavín, 2020; Jiménez-Sánchez et al., 2021).

The next factor is demographic. The arrival in European countries of emigrants from Maghreb, Sub-Saharan Africa and Latin America has also brought a new influx of players, especially second-generation emigrants, to the national championships Without emigration all these players would nurture the championships of their parents' countries of origin, but now they are present in the championships of the countries where they were born. This has led to an increase in the number of international players in Europe, as many of them either play for their country of birth or for the national team of their parent's country.

The next change is related to the previous one. More and more young players are signed from other countries, either to nurture the youth academies or to acquire them before their price becomes too high. (Najarian, 2015; Nicolls & Worsfold, 2016), such as the cases of Messi at F.C. Barcelona or Kum Agüero at Atlético de Madrid (year 2000 and 2006 respectively). And although FIFA has tried to stop this bleeding of the weaker championships, the fact is that the current generation of players corresponds to a large extent to these characteristics.

In relation to them, another type of players are the youth-team players or those, which regardless of their nationality, come from the reserve or base team. Investment in youth academies can bring a lot of economic return (Kulikova & Goshunova, 2014), though the use of this players by the first team does not have to guarantee a better performance. This is because, while they may be cheaper and bring a sense of identity to the team and the fans, they may also be players with less experience or less quality than the rest of the professionals who make up a top-level league. Therefore, this research will also focus on the relevance of this type of players.

The last factor to review purely concerns sports. Since the late eighties with F.C. Barcelona led by Johan Cruyff and the early nineties with Arrigo Sacchi's AC Milan, football theory has been shaping a mainstream that focuses on touch football, looking for players of high technical quality. However, the physical characteristics of strength and aggressiveness are not underestimated, but they are not sacralised as it happened in the Italian and English championships of the eighties. Thus, in contrast to the footballing traditions mentioned above, there is an evolution towards offensive football, which has increased since the early 21st century, as it is beginning to be seen that attacking tactics bring success (Konstadinidou & Tsigilis, 2005; Luhtanen et al., 2001). This quest for quality and attacking play crystallised in the years from 2008 to 2014, with the game played by the Spanish team, the German team at national level and F.C. Barcelona led by Josep Guardiola at club level.

Another of the issues that would define such attacking football would be scoring chances, as manifested by shots on goal, possession, corner kicks and goals (Abellán et al., 2019; Borrás & Sainz de Baranda, 2005; Fradua et al., 2016; González-Ródenas et al., 2015; Iván-Baragaño, 2022). Other parameters such as the number of passes are much more imprecise, since they can occur in their own half and have no real bearing on scoring chances.

Thus, in summary, the major European championships are the ones with the most international players, the ones with the biggest financial budgets, the ones that are more committed to attacking football and the ones that are less permissive with faults, compared with other national championships and other teams. All these circumstances influence the development of a football that is increasingly a completely globalised spectacle and it starting to become more homogenous, at least in Europe.

Therefore, the aim of this research is to know in a quantitative way those variables that determine the final points and goals per game of Europe football teams, as well as to analyse the possible linear regression models for each country and competition in the 2019-2020 season analysed. Based on the above, different hypotheses can be established. The first is to find that the budget is one of the factors most related to the points. (Rossi et al., 2013). Secondly, to demonstrate the importance of different types of players in achieving victories, whether they are foreign (Royuela & Gásquez, 2018), home-grown or from international teams.

The third hypothesis proposes to find an interrelation between the number of goals per match and those variables involved in the style of play, where offensive elements (shots, corners kicks or goal scoring chances) as well as ball possession would be the most relevant in predicting a higher number of goals per match. Finally, taking into account the theory of social facilitation (Zajonc, 1965), the last hypothesis tries to demonstrate a higher number of goals for home teams compared to away teams and, in turn, more goals with a home crowd than without when playing at home.

Methodology

Eight European leagues whose final matches of the 2019-2020 season were played without a crowd were analysed. The first and the second division German Bundesliga, the Santander League (first and second division of Spanish professional football), the Italian League in both categories (Series A and B) and the English Premier League.

Being the same season and a not a new, each league continued to be made up the same teams, with no change of squad and their objectives remained the same as when they stopped playing with the public. This occurred at the beginning of March 2020, with a break of approximately two months and resuming at the end of May and beginning of June in these countries (with some exceptions).

The units of analysis are all the games played by each team. Thus, if for example the German Bundesliga first division had 306 matches, this results in a total of 612 cases (18 teams played 34 matches), of which half were played at home and half away (17 on each pitch).

For the study, a table was created with each competition containing technical data, name of the league, team, match day and date. It also includes the dependent variables, goals and points obtained (3 won / 1 drawn / 0 lost) both for each match and at the end of the season; and the rest of variables, which can be classified into external or general variables specific to each team, and internal or specific variables to each match. So, for the first block (external factors) the data is applied to each club regardless of the match. These are the mean attendance at the stadium (not counting when there were no spectators), the budget or the value of the team at the end of the season, the number of foreigners per club (born in a different country to one they play in and regardless of whether they are nationalised or count as EU players), the percentage of players (nationals/foreigners) fielded, the mean experience of the players (number of matches on the category), the experience of youth players or players taken from the club's own academy (number of matches with the first team), international players (who are or were in the national team of their country) and the final fair play (yellow plus red cards).

In a second block (internal factors) would be other party-specific and non-totals variables. These are number of faults, sum of yellow and red cards per match (the latter count as two warnings), ball possession (%), shots (sum of shots on target, shots wide and blocked shots), corner kicks, clear chances, percentage of successful dribbles and tackles (actions taken to get the ball away from the opponent).

In the case of the first block, it is not possible to differentiate between matches played at home or away, or those played with or without crowd, as the date are the same for these four variations. Though, in the second case, it is feasible to take into account the pitch and public factors since the data comes from each match. Therefore, the first group of variables will relate to the final points, while the second group will relate to the goals scored in each match.

Table 1. Descriptive of each league I

Descriptive (Sums, means, s.d., %)	German Bundesliga 1	German Bundesliga 2	Spanish League 1	Spanish League 2
Teams	18	18	20	22
Total matches	612	612	760	922
Matches before (with crowd)	446 (72.9%)	450 (73.6%)	538 (70.8%)	680 (73.7%)
Matches after (without crowd)	166 (27.1%)	162 (26.4%)	222 (29.2%)	242 (26.3%)
Match day	34	34	38	42
Total points (s.d.)	47 (15)	45 (8)	51 (15)	55 (8)
Goals (s.d.)	1.6 (1.3)	1.4 (1.2)	1.2 (1.1)	1.1 (1)
Budget in million euro (s.d.)	236 (204)	18 (12)	258 (268)	14 (10)
Stadium attendance (s.d.)	40.940 (18.047)	20.211 (12.851)	29.136 (19.166)	9.870 (5.277)
No. of foreigners for team (%)	16 (51%)	6.5 (25%)	10.3 (38%)	5.8 (27.1%)
Regular players	Foreigners: 58% Nationals: 42%	Foreigners: 32% Nationals: 68%	Foreigners: 40% Nationals: 60%	Foreigners: 24% Nationals: 76%
Experience (s.d.) (Matches played per player)	65.2 (24)	51 (16)	91 (29)	82 (20)
Experience of the home-grown players (s.d.)	244 (183)	55.1 (41)	463.2 (466)	121.4 (138)
Internationals (s.d.)	6 (3)	.6 (.6)	4 (3)	1 (1)
Fair play (s.d.) (total cards)	73 (13)	78 (14)	107 (21)	128 (15)
Shots (s.d.)	13.2 (5)	13.5 (4.7)	11.3 (4.5)	11.2 (4.3)
Possession at home (s.d.)	51.4% (12)	50.8% (11)	50.9% (11)	50.9% (9)
Possession away(s.d.)	48.6% (12)	49.2% (11)	49.1% (11)	49.1% (9)
Corner kicks (s.d.)	5 (3)	5.1 (2.7)	4.6 (2.5)	4.4 (2.5)
Faults (s.d.)	12.1 (4)	13 (4)	13.7 (4.1)	15.7 (4.2)
Sum of cards (s.d.)	2.1 (1.5)	2.2 (1.5)	2.7 (1.7)	3 (1.6)
Chances (s.d.)	2.4 (1.8)	2.1 (1.5)	1.8 (1.5)	1.5 (1.3)
Tackles (s.d.)	16.3 (5.1)	15 (5.1)	15 (4.5)	14.6 (4.6)
% Successful dribbles (s.d.)	59% (13)	58% (14)	59% (14)	58% (14)

Note: s.d. = Standard Deviation

Table 2. Descriptive of each league II

Descriptive (Sums, means, s.d., %)	Series A Italy 1	Series B Italy 2	Premier League England	Bundesliga Austria
Teams	20	20	20	12
Total matches	760	760	760	384
Matches before (with crowd)	499 (65.7%)	528 (69.5%)	576 (75.8)	264 (68.8%)
Matches after (without crowd)	261 (34.3%)	232 (30.5%)	184 (24.2)	120 (31.2%)
Match day	38	38	38	32 (22+10)
Total points (s.d.)	52 (17)	51 (12)	52 (17)	44 (14)
Goals (s.d.)	1.5 (1.2)	1.2 (1)	1.3 (1.2)	1.6 (1.5)
Budget in million euro (s.d.)	229 (172)	15 (5)	415 (287)	20 (25)
Stadium attendance (s.d.)	27.195 (13.950)	6.135 (2.492)	39.366 (16.027)	6.687 (4.622)
No. of foreigners for teams (%)	17 (61.2%)	7.4 (26.1%)	18.3 (62.8%)	8 (29%)
Regular players	Foreigners: 62% Nationals: 38%	Foreigners: 22% Nationals: 78%	Foreigners: 64% Nationals: 36%	Foreigners: 30% Nationals: 70%
Experience (s.d.) (Matches played per player)	88.3 (23)	80 (19)	72.7 (30)	72 (16)
Experience of the home-grown players (s.d.)	136.1 (122)	52 (83)	201.5 (161)	147.5 (169)
Internationals (s.d.)	7 (4)	.6 (1)	8 (3)	2 (2)
Fair play (s.d.) (total cards)	109 (15)	113 (18)	72 (14)	78 (18)
Shots (s.d.)	14.3 (5.6)	13.2 (4.8)	12.3 (5.4)	12.8 (5.5)
Possession at home(s.d.)	50.1% (10)	49.8% (9)	50.1% (12)	50.8% (10)
Possession away (s.d.)	49.9% (10)	50.2% (9)	49.9% (12)	49.2% (10)
Corner kicks (s.d.)	5.3 (3)	4.7 (2.4)	5.3 (3)	5.1 (2.9)
Faults (s.d.)	13.9 (4)	16 (4.3)	10.7 (3.6)	15.2 (4.2)
Sum of cards (s.d.)	2.9 (1.7)	2.8 (1.6)	1.8 (1.3)	2.5 (1.6)
Chances (s.d.)	2 (1.6)	1.8 (1.4)	2 (1.6)	2.2 (1.8)
Tackles (s.d.)	14.5 (4.5)	14 (4.4)	16.3 (5.2)	15.7 (5.1)
% Successful dribbles (s.d.)	61% (14)	56% (15)	60% (14)	57% (16)

Note: s.d. = Standard Deviation

All the data were collected from the following sources:

- https://sofascore.com
- www.transfermarkt.es
- https://es.soccerway.com/

These websites are of recognised prestige and responded correctly to the data quality test conducted on ten random matches from each league.

Once the data had been collected, the statistical programme SPSS (version 23 for Windows) was used. The tests used were Pearson's correlation, step-by-step multiple linear regression by the intro method, taking total points and goals per match as dependent variables, and the rest of variables mentioned as independent variables.

Results

Before presenting the main analyses, it is necessary to mention a series of fundamental correlations (Pearson, $p < .05$). Thus, across the board and bidirectionally, more points at the end of the season and goals scored per game correlate with higher team Budget, higher mean stadium attendance, number of foreign players and use of them (fielded), more fair play and fewer cards per match, more experience of the players (in that category), less track record in the case of home-grown players, more ball possession, shots and goal scoring chances (see Appendices).

The analyses carried out are shown below, recalling that the increase in models entails the inclusion of the variables corresponding to the preceding ones. The first row shows the variables that each model adds to the previous one, in brackets the adjusted R^2 corresponding to the step-by-step final model. The middle row shows the contribution of that variable to the dependent variable in an individualised model with the intro method, except in the first of them, where the adjusted R^2 would be the same as the general one because it is the first predict. Finally, in the third row the standardised beta with its p-value.

Table 3. Lineal regression for the first block of variables I. Dependent = Total points.

Model	**Championships** **Variable (adjusted *R*2 in the step-by-step final model)** ***Individual adjusted R*2 with DV** **Standardised Beta (sig)**			
	Bundesliga 1	Bundesliga 2	Spanish League 1	Spanish League 2
1	Budget (.824) β=.908 (.000)	Fair play (.432) β=-.658 (.000)	Internationals (.665) β=.816 (.000)	Budget (.467) β=.663 (.000)
2	Experience (.859) *Ad. R^2*=.620 β=.788 (.000)	Budget (.578) *Ad. R^2*=.263 β=.514 (.000)	Fair play (.720) *Ad. R^2*=.115 β=-.341 (.000)	Foreigners fielded (.491) *Ad. R^2*=.021 β=.150 (.000)
3	Attendance (.882) *Ad. R^2*=.268 β=.518 (.000)	Experience (.662) *Ad. R^2*=.111 β=.336 (.000)	Budget (.745) *Ad. R^2*=.658 β=.811 (.000)	Fair play (.502) *Ad. R^2*=.020 β=.146 (.000)
4	Fair play (.888) *Ad. R^2*=.399 β=-.632 (.000)	Attendance (.716) *Ad. R^2*=.102 β=.322 (.000)	Foreigners (.748) *Ad. R^2*=.365 β=.605 (.000)	Attendance (.508) *Ad. R^2*=.003 β=.062 (.058)
5	Internationals (.891) *Ad. R^2*=.648 β=.805 (.000)	Internationals (.739) *Ad. R^2*=.220 β=.471 (.000)		
6	Foreigners (.896) *Ad. R^2*=.061 β=.251 (.000)	Foreigners (.746) *Ad. R^2*=.049 β=.225 (.000)		
7	Foreigners fielded (.905) *Ad. R^2*=.052 β=.231 (.000)	Home-grown (.760) *Ad. R^2*=.001 β=.025 (.538)		
8	Home-grown (.916) *Ad. R^2*=.229 β=.480 (.000)	Foreigners fielded (.771) *Ad. R^2*=.028 β=.173 (.000)		

Table 4. Lineal regression for the first block of variables II. Dependent = Total points

Model	**Championships Variable (adjusted R^2 in step-by-step final model) *Individual adjusted* R^2 with DV. Standardised Beta (sig)**			
	Series A	Series B	Premier	Austria
1	Budget (.697) β=.835 (.000)	Internationals (.321) β=.542 (.000)	Budget (.712) β=.844 (.000)	Fair play (.589) β=-.768 (.000)
2	Foreigners fielded (.768) *Ad.* R^2=.520 β=.722 (.000)	Foreigners fielded (.411) *Ad.* R^2=.000 β=.039 (.277)	Fair play (.797) *Ad.* R^2=.177 β=-.422 (.000)	International (.740) *Ad.* R^2=.454 β=.675 (.000)
3	International (.801) *Ad.* R^2=.451 β=.672 (.000)	Experience (.499) *Ad.* R^2=.042 β=-.209 (.000)	Foreigners (.860) *Ad.* R^2=.053 β=.234 (.000)	Attendance (.755) *Ad.* R^2=.319 β=.566 (.000)
4	Attendance (.807) *Ad.* R^2=.486 β=.698 (.000)	Fair play (.594) *Ad.* R^2=.067 β=-.262 (.000)	Home-grown (.874) *Ad.* R^2=.006 β=.084 (.020)	
5	Foreigners (.818) *Ad.* R^2=.263 β=.514 (.000)	Home-grown (.597) *Ad.* R^2=.010 β=-.108 (.004)	Foreigners fielded (.878) *Ad.* R^2=.058 β=.243 (.000)	
6	Home-grown (.826) *Ad.* R^2=.063 β=.253 (.000)		Experience (.879) *Ad.* R^2=.080 β=.286 (.000)	
7	Experience (.827) *Ad.* R^2=.309 β=.556 (.000)		Attendance (.879) *Ad.* R^2=.234 β=.484 (.000)	

Then, for the second bloc, the variables pitch (home / away) and crowd were added.

Table 5. Lineal regression for the second block of variables I. Dependent = Goals scored per match

Model	**Championships Variable (adjusted R^2 step-by-step final model) Individual adjusted R2 with DV. Standardised (sig)**			
	Bundesliga 1	Bundesliga 2	Spanish League 1	Spanish League 2
1	Chances (.379) β=.617 (.000)	Chances (.281) β=.532 (.000)	Chances (.308) β=.556 (.000)	Chances (.277) β=.527 (.000)
2	Cards (.387) *Ad.* R^2=.019 β=-.142 (.000)	Possession (.290) *Ad.* R^2=-.001 β=-.028 (.493)	Corner kicks (.327) *Ad.* R^2=.002 β=-.054 (.137)	Possession (.297) *Ad.* R^2=.019 β=-.140 (.000)
3	Corner kicks (.395) *Ad.* R^2=-.002 β=-.011 (.777)	Corner kicks (.293) *Ad.* R^2=.002 β=-.057 (.160)	Pitch (.340) *Ad.* R^2=.029 β=-.174 (.000)	Pitch (.300) *Ad.* R^2=.016 β=-.129 (.000)
4	Shots (.403) *Ad.* R^2=.081 β=.287 (.000)	Shots (.301) *Ad.* R^2=.037 β=.197 (.000)	Shots (.347) *Ad.* R^2=.058 β=.244 (.000)	Dribbles (.303) *Ad.* R^2=.006 β=.086 (.009)
5	Possession (.406) *Ad.* R^2=.037 β=.197 (.000)		Cards (.353) *Ad.* R^2=.016 β=-.133 (.000)	Tackles (.305) *Ad.* R^2=.000 β=-.032 (.331)

In summary, the multiple results confirm the importance of budget, fair play, scoring chances and possession in achieving victories, while questioning the effectiveness of the presence of the crowd.

Table 6. Lineal regression for the second block of variables II. Dependent = Goals scored per match

Model	**Championships** Variable (adjusted R^2 in the step-by-step final model) *Individual adjusted* R^2 with DV. Standardised Beta (sig)			
	Series A	Series B	Premier	Austria
1	Chances (.290) β=.540 (.000)	Chances (.243) β=.494 (.000)	Chances (.343) β=.586 (.000)	Chances (.446) β=.669 (.000)
2	Corner kicks (.297) *Ad.* R^2=-.001 β=-.013 (.717)	Shots (.257) *Ad.* R^2=.091 β=.303 (.000)	Shots (.354) *Ad.* R^2=.145 β=.383 (.000)	Shots (.458) *Ad.* R^2=.215 β=.466 (.000)
3	Shots (.309) *Ad.* R^2=.069 β=.264 (.000)	Possessions (.281) *Ad.* R^2=.012 β=-.113 (.002)	Corner kicks (.367) *Ad.* R^2=.005 β=.078 (.032)	Cards (.464) *Ad.* R^2=.059 β=-.248 (.000)
4	Crowd (.312) *Ad.* R^2=.002 β=.057 (.118)	Corner kicks (.285) *Ad.* R^2=-.001 β=.004 (.916)	Pitch (.372) *Ad.* R^2=.015 β=-.126 (.001)	
5		Pitch (.289) *Ad.* R^2=.018 β=-.140 (.000)	Cards (.375) *Ad.* R^2=.017 β=-.134 (.000)	

Discussion

Regarding the first block of variables analysed, the first established hypothesis is confirmed, as it is generally observed that the budget is what most influences the achievement of victories (Rossi et al., 2013). Obviously, this monetary factor would also determine the value of other variables (attendance, signing foreigners, and more experienced and international players). However, it should be borne in mind that it is not homogeneous and that are differences and exceptions between the competitions studied.

Thereby, the budget is found to be a very significant element in determining the total points, only in Series B and in the Austrian League, while it is not present in the final regression. This pattern coincides with the budget amount seen above (Table 1 and Table 2), as it is most influential in the leagues with the largest imbalances (standard deviations). However, in competitions with teams more equal in value this factor would not be so decisive. This is observed in the Austrian league, where the monetary difference is smaller and the budget is not present in the final linear regression model. Therefore, it is not so much the amount of the club's budget that influences the final score but its relation to the values of the other teams in that league (more equal, less influence).

Obviously, the budget amount alone is not the cause of the number of points or a better game, but it is the one that makes possible the other elements that are directly correlated to the achievement of victories. Although, as has been analysed, each championship would have different investment proposals to achieve them, be it for example by signing foreigners in the case of the Italian leagues, or by advocating for less aggressive players and with more fair play, as in the German second league or the Austrian Bundesliga. In other words, depending on the type of championship and country, the budget should be invested in one type of player or another (Magaz-González et al., 2017).

However, it should be mentioned that the relationship between budget and goals is not so clear. It has already been shown in Jiménez–Sánchez & Lavín (2020) that, in the matchdays without crowd due to the anti-Covid measures, the points obtained by teams were correlated with the budget, but not so much with the number of goals. Therefore, it can be extrapolated that scoring in a match depends on a number of factors that have less to do with money.

In reference to the second hypothesis, this is clarified with the data showing the importance of the different types of players in achieving victories. In this sense, it is observed that the volume of international players or those who have played for their country's national team is present in the final model of all leagues except the Spanish second league and the Premier League. Moreover, the standardised beta reflects that this factor is quite decisive in all competitions (p <.001), and even in Series B more so than budget, which does not even correlate significantly with points.

In turn, the use of foreigners is also decisive in the final models of all competitions except the Austrian one, although individually it is significant for the points (p <.000). The correlations reflect a higher incidence in the Spanish and Italian first division.

Additionally, the experience of players and home-grown players is present in the final model of most competitions except the both Spanish leagues and in the Austrian. Although they do correlate individually, and even the Spanish and Italian second positions correlate inversely, that is, the more experienced the players are, the fewer points the team scores and vice versa.

In summary, it is generally observed that foreigners and those who play or have played in international teams contribute more to the achievement of points than home-grown players and even experienced professionals. These results may be useful to support transfer investment strategies, although it is worth mentioning that the categories proposed can be improved for further studies, for example, by excluding nationalized or EU citizens as foreigners, or taken into account the volume of games played in the case of football players in international teams. Furthermore, it must be considered that these categories are not mutually exclusive and can easily overlap, that is, players like Lionel Messi, according to Transfermarkt, count as foreign international, home-grown and experienced at the same time.

Furthermore, the third hypothesis is partially confirmed by the data revealing that chances are the most decisive factor in the number of goals scored, which is quite logical, as more chances mean more chances to score. The Austrian Bundesliga is the one with the most correlation (β = .669) and Series B the lowest (β = .494).

At the same time, it shows how shots and corner kicks (both elements of offensiveness) are present in the regression models of most competitions except in the Spanish second division. The difference between the two is that the shots correlate much more in all competitions with respect to corner kicks.

Regarding possession, contrary to what might be expected, it is only found in four competitions, and in some cases the correlations are even inverse (greater possession fewer goals) as in the second divisions of Germany, Spain and Italy. Thus, possession of the ball, present in final regression models of half of the championships analysed, should not be valued as a maxim to be achieved or assume that possession alone benefits goals per game, as it has been generally conceived at a historical and tactical level (Lago & Martín, 2007). But, it is rather a complement that can favour the other more decisive components such as the creation of scoring chances or shots on goal.

Concerning cautions it was also found in the final model in four competitions, while the number of faults is not present in any model. On both variables the correlations have the same tendency as the one mentioned above fair play (fewer faults and cautions, more goals scored).

As for the tackling and dribbling success rates, they are only significant decisive in the Spanish second division. For the most part, the lower the number of tackles, the higher the number of goals and the positive correlation in the case of dribbling.

It should be recalled that the fact that certain variables are not part of the model does not mean that they are not individually correlated. Accordingly, the final model in the Austrian league only takes into account three of the total number of variables introduced. In contrast, the Pearson correlation ($p < .001$) of each of these with goals, in addition to correlating with these three, also correlates with possession and corner kicks, but effectively with a lower and insufficient relevance to form part of the final resulting model in that country.

In reference to the last hypothesis, this is also partially confirmed, because although the home team scores more goals in most of the leagues analysed, this is not due to the attendance of spectators. In this sense, the lower importance of the crowd than expected is striking, as it is only included in the linear regression model of Series A and also with an insignificant p-value (.118). However, as shown in tables 3 and 4, the number of stadium attendees does appear as a relevant factor in the achievement of points. These results would reveal that it is not so much the volume of fans that influences the goals, but that the best teams (most point) are those that also happen to be the ones with the highest stadium attendance. But that does not mean that the presence of fans helps to score goals, in other words, the absence of a crowd does not affect the game as significantly as one might expected.

To clarify this, it should be recalled that stadium attendance is present in the models obtained in most competitions analysed, except in the Spanish first division and the Italian second division. The individualised model again shows an association with the differences between clubs. That is, the more discrepancy there is in the average attendance of the teams, the higher the standardised beta and the higher the adjusted R^2. And the more equal the attendance between their teams, the less it contributes to the regression model, as it would be cancelled out by tending towards a constant. Since stadium capacity and attendance have two-way relationship with the team's budget (more money allows for a bigger stadium and higher attendance generates more revenue), the results found make it possible to equate the number of fans at the same level as the budget, as if it were a single factor under the same operating pattern.

Therefore, although field factor, with more goals at home tan away, is registered in both Spanish competitions, the Italian second division and the English one, contrary to what is proposed by theory of social facilitation (Zajonc, 1965) or the influence of spectators on home victories (Pollard, 2008; Pollard y Gómez, 2014), none of this four models contains the crowd factor. The only league that does have it is Series A and, furthermore, with a high number of goals without fans that with them for both home and away matches, the opposite that what was theoretically proposed

It is worth mentioning the role of fair play, since this is shown as a very relevant factor, less so in the final model of the Italian first division, although individually is very much related ($r = -.202$; $p < .000$). Regarding this aspect, it can be deduced that a lower number of cards at the end of the season leads to more points, as these teams would have been those less prone to opposing attacks or with less aggressive defences, as well as having fewer players sent off. However, the Spanish second division shows the opposite, less fair play more points, although it is worth mentioning that its R^2 is quite low and would contribute almost nothing to the dependent variable. Therefore, the increase of cards seems to be an element to avoid in general in the pursuit of victories. Thus, fair play should be a maxim to be followed by all teams to demonstrate not only that fair play is feasible and favourable, but also to offer and promote a healthy game to the fans where infringements are detrimental for their teams to win.

Regarding the limitations, it should be remembered that the 2019-2020 season has been selected for this research, the last matches of which were played without spectators, an aspect that may have had a particular influence on the internal factors. Therefore, it is recommended to study other old seasons in order to compare them with this one and thus observe whether the public effect has led to changes in the components of the game, while at the same time this comparison would serve to analyse the evolution of each league investigated. It is also advisable for future studies to extend the sample with non-European championships and even take women's league as a reference to see how these variables behave. In addition, the same regression models can be carried out for other sports with the same or more media coverage in other countries, such as basketball or rugby.

Conclusions

This research has sought to find out how important certain factors are in scoring points and goals in eight European leagues in the 2019-2020 season in the 2019-2020 season. As stated at the beginning, the results allow us conclude that the budget is the most relevant element in most of the leagues studied. This occurs especially in those with three-digit budget averages where there would also be a large economic difference between their teams.

This variable would be closely associated with the rest of the components in a bidirectional way, as a larger budget would condition, for example, the size of the capacity or the signing of better players, which would be associated with a better style of play, with more shots and goal scoring opportunities among others. This would lead to more goals and points, that is, a better league position, which would have an impact on the economic benefit for the following season. In this way, a spiral is formed that gradually tends to benefit those who have the most, but also progressively harm those with the least money. It is because of this same relationship between these variables that we would find differences between individual leagues. Some would be more equal in terms of budget and this would not be so decisive in the final ranking, while other leagues would be less equal financially and where it would be easier to predict the achievements at the end of season.

Finally, the practical implications are, on the one hand, to take advantage of the analysis of external factors by the various clubs to improve their investment strategies. On the other hand, the results and methodology used for internal factors can be useful to determine precisely the different styles of play of each team and thus prepare for opponents with specific training and tactics (Prieto-Ayuso et al., 2022; Serra-Olivares et al., 2021).

In short, it can be seen that the first block has more robust predictive models than in the cases of the second block analysed. Therefore, it would be easier to predict points at the end of the season than goals per game. In any case, it is necessary to take into account the similarities as well as the specific and differential elements of each league and to avoid generalisations as far as possible. Likewise, the evolution of the contributions made by the variables analysed both in terms of points as well as goals should be studied longitudinally.

Primary data availability

Pearson correlations between variables by country: https://uvadoc.uva.es/handle/10324/59194

Bibliography

Abellán, J., Sáez-Gallego, N., Savelsbergh, G., & Contreras, O. (2019). La intercepción de un lanzamiento de córner desde la teoría de los limitadores. *Revista Internacional De Medicina Y Ciencias De La Actividad Física Y Del Deporte*, *19*(74). https://doi:10.15366/rimcafd2019.74.004

Andrikopoulos, A., & Kaimenakis, N. (2006). Introducing FOrNeX: A composite index for the intangible resources of the football club. *International Journal of Sport Management* and Marketing, *5*(3), 251-266. https://doi:10.2139/ssrn.907675

Borrás, D., & Sainz de Baranda, P. (2013). Análisis del córner en función del momento del partido en el mundial de Corea y Japón 2002. *Cultura, Ciencia y Deporte, 1*(2), 87-93. https://doi:10.12800/ccd.v1i2.90

Bullough, S., & Jordan, J. (2017). Youth academy player development in English football: The impact of regulation since 2006. *Sport, Business and Management, 7*(4), 375-392. https://doi:10.1108/SBM-10-2016-0059

Carolina-Paludo, A., Nunes-Rabelo, F., Maciel-Batista, M., Rúbila-Maciel, I., Peikriszwili-Tartaruga, M., & Simões, C. A. (2020). Game Location Effect on Pre-competition Cortisol Concentration and Anxiety State: A Case Study in a Futsal Team. *Revista De Psicología Del Deporte*, *29*(1), 105–112. https://www.rpd-online.com/index.php/rpd/article/view/40

Fothergill, M., Wolfson, S., & Neave, N. (2017). Testosterone and cortisol responses in male soccer players: The effect of home and away venues. *Physiology & Behavior 177*(1), 215-220. https://doi:10.1016/j.physbeh.2017.04.021

Fradua, L., Zubillaga, A., & Ford, P R. (2016). Attacking and defensive styles of play in soccer: analysis of Spanish and English elite teams. *Journal of Sports Sciences, 34*(24), 2195-2204. https://doi:10.1080/02640414.2016.1169309

García-Vega, M.A. (10 de Enero, 2020). Una burbuja llamada fútbol. *Diario El País*. https://elpais.com/economia/2020/01/16/actualidad/1579196525_489238.html

González-Ródenas, J., López-Buendía, I., Calabuig-Moreno, F., & Aranda-Malavés, R. (2015). Indicadores tácticos asociados a la creación de ocasiones de gol en fútbol profesional. *Cultura, Ciencia y Deporte, 10*(30), 215-225. https://doi:10.12800/ccd.v10i30.590

Hamil, S., Walters, G., & Watson, L. (2010). The model of governance at FC Barcelona: balancing member democracy, commercial strategy, corporate social responsibility and sporting performance. *Soccer & Society, 11*(4), 475-504. https://doi:10.1080/14660971003780446

Hughes, M., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of Sport Sciences, 23*(5), 509-514. https://doi:10.1080/02640410410001716779

Iván-Baragaño, I., Maneiro, R., Losada, J. L. & Ardá, A. (2022). Tactical Differences Between Winning and Losing Teams in Elite Women's Football. *Apunts Educación Física y Deportes, 1*(147), 45-54. https://doi:10.5672/apunts.2014-0983.es.(2022/1).147.05

Jiménez-Sánchez, Á. & Lavín, J.M. (2020) Home advantage in European soccer without crowd, *Soccer & Society*. https://doi:10.1080/14660970.2020.1830067

Jiménez-Sánchez, Á, Lavín, J. M., & Endara, F. (2021). Repercusiones de jugar sin público en la ventaja local, las decisiones arbitrales y en los componentes del juego. *Cuadernos de Psicología del Deporte*, *21*(2), 198–212. https://doi:10.6018/cpd.444811

Katz, M., Thomas A. Baker, T.A., & Hui Du. (2020). Team Identity, Supporter Club Identity, and Fan Relationships: A Brand Community Network Analysis of a Soccer Supporters Club. *Journal of Sport Management, 34*(1): 9–21. https://doi:10.1123/jsm.2018-0344

Konstadinidou, X. & Tsigilis, N. (2005) Offensive playing profiles of football teams from the 1999 Women's World Cup Finals. *International Journal of Performance Analysis in Sport*, *5*(1), pp 61-71. https://doi:10.1080/24748668.2005.11868316

Kulikova, L. I., & Goshunova, A. V. (2014). Human capital accounting in professional sport: evidence from youth professional football. *Mediterranean Journal of Social Sciences, 5*(24), 44. https://doi:10.5901/mjss.2014.v5n24p44

Lago, C., & Martín, R. (2007). Determinants of possession of the ball in soccer. *Journal of Sports Sciences, 25*(9), 969-974. https://doi:10.1080/02640410600944626

Lago-Ballesteros, J., Lago-Peñas, C., & Rey, E. (2012) The effect of playing tactics and situational variables on achieving score-box possessions in a professional soccer team. *Journal of Sports Sciences, 30*(14), 1455-1461. https://doi:10.1080/02640414.2012.712715

López-Busto, A., García-Unanue, J., Gómez-González, C., Barajas, Á., & Gallardo Guerrero, L. (2016). Incidencia de los resultados deportivos, las variables económicas y administrativas en el rendimiento financiero de los clubes de fútbol. *Cultura, Ciencia y Deporte, 11*(33), 185-195. https://doi:10.12800/ccd.v11i33.763

Luhtanen, P, Belinskij, A., Häyrinen, M. & Vänttinen, T. (2001). A comparative tournament analysis between the EURO 1996 and 2000 in soccer. *International Journal of Performance Analysis, 1*, 77-83. https://doi:10.1080/24748668.2001.11868250

Magaz-González, A., Mallo-Fernández, F., & Fanjul-Suárez, J. (2017). ¿Es rentable jugar en Primera División de fútbol? / Is profitable to play in Spanish Soccer First Division?. *Revista Internacional De Medicina Y Ciencias De La Actividad Física Y Del Deporte*, (65). https://doi:10.15366/rimcafd2017.65.001

Najarian, A. (2015). The Lost Boys: FIFA's Insufficient Efforts to Stop Trafficking of Youth Footballers. *Sports Law Journal 22,* 151–192.

Neave, N., & Wolfson, S. (2003). Testosterone, territoriality, and the 'home advantage'. *Physiology & Behavior, 78*(2), 269-275. https://doi:10.1016/S0031-9384(02)00969-1

Nicolls, S., & Worsfold, P. R. (2016). The observational analysis of elite coaches within youth soccer: The importance of performance analysis. *International Journal of Sports Science and Coaching, 11*(6), 825-831. https://doi:10.1177/1747954116676109

Peña, J.L., & Navarro, R. (2015). Who can replace Xavi? A passing motif analysis of football players. *ArXiv,* abs/1506.07768. https://doi:10.48550/arXiv.1506.07768

Poli, R., Ravenel, L., & Besson, R. (2016). Foreign players in football teams. *CIES Football Observatory Monthly Report,* (12), 1-9. https://football-observatory.com/IMG/pdf/mr12_eng.pdf

Pollard, R. (2008). Home advantage in soccer: A retrospective analysis. *Journal of Sports Sciences, 4*(3), 237-248. https://doi:10.1080/02640418608732122

Pollard, R., & Gómez, M. A. (2014). Components of home advantage in 157 national soccer leagues worldwide. *International Journal of Sport and Exercise Psychology, 12*(3), 218-233. https://doi:10.1080/1612197X.2014.888245

Prieto-Ayuso, A., Pastor-Vicedo, J. C., Martínez-Martínez, J., & Contreras-Jordán, O. R. (2022). Tactical Performance of Talented Youth Soccer Players. *Cultura, Ciencia y Deporte, 17*(51), 145-152. https://doi:10.12800/ccd.v17i51.1721

Rossi, M., Thrassou, A., & Vrontis, D. (2013). Football performance and strategic choices in Italy and beyond. *International Journal of Organizational Analysis*, *21*(4), 546–564. https://doi:10.1108/IJOA-04-2013-0659

Royuela, V., & Gásquez, R. (2018). On the influence of foreign players on the success of football clubs. *Journal of Sports Economics, 20*(5), 718-741. https://doi:10.1177/1527002518807960

Sarmento, H., Anguera, M.T., Campaniço, J., Resende, R. & Leitão, J. (2014). Manchester United, Internazionale Milano and FC Barcelona – What´s different? *Sport Mont Journal, 12*(40-41-42), 50-56. http://www.sportmont.ucg.ac.me/?sekcija=article&artid=173

Serra-Olivares, J., Maleki, M., & Zarei, A. (2021). Role of Age and Experience in Attacking Tactical Knowledge among Iranian Soccer School Players. *Cultura, Ciencia y Deporte, 16*(49), 485-496. https://doi:10.12800/ccd.v16i49.1453

Solberg, H. A. & Mehus, I. (2014). The challenge of attracting football fans to stadia? *International Journal of Sport Finance, 9*(1), 3–19. https://www.academia.edu/17717157/The_Challenge_of_Attracting_Football_Fans_to_Stadia

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*(3681), 269-274. https://doi:10.1126/science.149.3681.269