



# Impact of Scoring First on Match Outcome in the Chinese Football Super League

Tianbiao Liu<sup>1,2</sup>, Antonio García-de-Alcaraz<sup>3,4\*</sup>, Hai Wang<sup>1,2</sup>, Ping Hu<sup>5</sup> and Qiu Chen<sup>6\*</sup>

<sup>1</sup>College of P.E. and Sports, Beijing Normal University, Beijing, China, <sup>2</sup>Lab of Performance Analysis and Motor Learning, Beijing Normal University, Beijing, China, <sup>3</sup>Department of Education, University of Almería, Almería, Spain, <sup>4</sup>LFE Research Group, Universidad Politécnica de Madrid, Madrid, Spain, <sup>5</sup>Microsoft, Beijing, China, <sup>6</sup>National School of Development, Peking University, Beijing, China

## OPEN ACCESS

### Edited by:

Nimai Parmar,  
Middlesex University, United Kingdom

### Reviewed by:

Miguel-Angel Gomez-Ruano,  
Polytechnic University of Madrid,  
Spain

John William Francis,  
University of Worcester,  
United Kingdom

### \*Correspondence:

Antonio García-de-Alcaraz  
antonogadealse@gmail.com  
Qiu Chen  
chenqiusau@163.com

### Specialty section:

This article was submitted to  
Performance Science,  
a section of the journal  
Frontiers in Psychology

Received: 01 February 2021

Accepted: 13 April 2021

Published: 28 May 2021

### Citation:

Liu T, García-de-Alcaraz A, Wang H,  
Hu P and Chen Q (2021) Impact of  
Scoring First on Match Outcome in  
the Chinese Football Super League.  
*Front. Psychol.* 12:662708.  
doi: 10.3389/fpsyg.2021.662708

The aim of this study was to explore the effect of scoring first on match outcomes in the Chinese Football Super League (CSL). A total of 1,116 matches in which at least one goal was scored from the 2014 to 2018 seasons were collected. Match outcomes, absolute goal differences, the minute of the first goal, match locations, and teams' budgets were analyzed. A team's budget was measured in terms of a team's value at the beginning of the season, and teams were clustered into two groups (high and low budget with means of 50.77 and 13.77 million dollars, respectively). A descriptive analysis was conducted, and two generalized linear models (a multinomial logit model and a Poisson model;  $p < 0.05$ ) were applied. The results showed a favorable outcome for the team that scored first both in match outcome and goal difference. Regarding the teams that scored first, 66.31% won their matches, 20.70% achieved a draw, and 12.99% lost. Specifically, home teams were more likely to win (13.42%) and less likely to lose (9.52%) or draw (3.90%) than away teams. Home teams also had a higher likelihood of obtaining a larger goal difference. Higher budget teams were more likely to win (14.90%) and less likely to lose (9.75%) or draw (5.14%) than low-budget teams. Additionally, for each minute, the team scores closer to the end of the match, and the average probability of winning increased by 0.0028. These findings can guide the strategies of coaches in different match scenarios according to the match location and the opponent's quality.

**Keywords:** team sports, performance analysis, score, match status, situational variables, goal

## INTRODUCTION

Football is a time-dependent sport whose result depends on comparing the goals of the two participating teams at the end of the match (Hughes and Franks, 2004). However, most matches ended with no more than three goals (Anderson and Sally, 2013). Therefore, every goal is critical and influences the course of the game (Ferreira, 2013). In many team sports, scoring first has already been proven to be crucial to the match outcome (Jones, 2009). Moreover, when the match approaches the end, the result is nearly determined (Altarriba-Bartés et al., 2020). Overall, scoring first has a strong influence on match outcome (Bloomfield et al., 2005). Across worldwide football tournaments, scoring first has been shown to be an effective predictor

of the final outcome in the 2018 FIFA (Fédération Internationale de Football Association) World Cup (71.40%; Vergonis et al., 2019), the six major football tournaments in the world (71.17%; Leite, 2015), the Greek league in the 2006–2007 (71.43%; Armatas et al., 2009b) and 2007–2008 seasons (74.20%; Armatas et al., 2009a), and the La Liga (Spanish League) from the 2005 to 2010 seasons (79.32%; Molinuevo and Bermejo, 2012).

The team that scores first breaks the competitive balance and forms a new match status. This shift of match status may lead to changes in match strategies, tactics, and styles of play for each team (Taylor et al., 2008; Lago, 2009; Almeida et al., 2014) and may even alter players' psychological states and, hence, influence their performance (Bar-Eli et al., 2006; Leite and Barreira, 2014). For instance, several studies have found a greater ball possession time among losing teams (Lago-Peñas and Dellal, 2010; Taylor et al., 2010; Ruiz-Ruiz et al., 2013). Although scoring first appears to be a key event in winning a match, sometimes the team who scores first loses. The complex interaction between multiple factors is a plausible reason that should be investigated (Garganta, 1997). In this sense, the situational variables have been considered. Thus, match location has been considered to be a strong indicator (Molinuevo and Bermejo, 2012; Tenga, 2013; Pratas et al., 2016). Home teams have better technical and tactical performance (Sasaki et al., 1999; Tucker et al., 2005), especially in attacking-related indicators. In the CSL, home field advantage is also reported (Liu et al., 2019a; Zhou et al., 2019), although its long-term impact fluctuates (Zhou et al. 2020). However, the quality of the opposition is considered a stronger determinant than match location (Liu et al., 2016). Linked to this team's ability, team budget is an indicator used to measure the strength of a team. Team budget has been proven to affect teams' performance (Lago-Peñas and Gomez-Lopez, 2014; Lago-Peñas and Sampaio, 2015; Liu et al., 2019b) as richer teams cannot only afford the best coaches and players but also provide the best training and playing conditions and then adapt differently to new season performance (Lago-Peñas and Gomez-Lopez, 2014). Hence, the higher the budget the teams scoring first have, the greater their winning probabilities (Gómez-Ruano et al., 2013). Moreover, the influence of the first scoring time or match period on match performance (García-Rubio et al., 2015; Lago-Peñas et al., 2016; Ibáñez et al., 2018; Martínez and Gonzalez-García, 2019) was also studied, but these results are still contradictory and need further investigation.

In modern sports, statistical and machine learning models are increasingly incorporated in the analysis of football games, which helps both researchers and practitioners better and more deeply understand game performance and its important influencing events (e.g., goals, scoring, fouls, etc.) and, hence, change the path to success. In particular, regression models (García-Rubio et al., 2015), classification trees (Lago-Peñas et al., 2016) and Cox models (Nevo and Ritov, 2013) were used to evaluate the effect of scoring first on match outcomes. Compared with early studies focusing on frequency, descriptive analysis, and comparisons (Yiannakos and Armatas, 2006; Armatas et al., 2009a; Leite, 2015), machine learning models can handle larger datasets and nonlinear relationships and then offer more

knowledge to researchers (Hucaljuk and Rakipović, 2011; Liu, 2014; Liu et al., 2018).

Studies on match analysis in the Chinese Football Super League (CSL) have received considerable attention starting a few years ago because of its fast progress. These research studies focus primarily on key performance indicators (Mao et al., 2016; Yang et al., 2018; Gai et al., 2019), the impact of situational variables (Liu et al., 2019b), the evaluation of performance (Zhou et al., 2018, 2020), and the application of new technologies (Han et al., 2020) and complex networks (Yu et al., 2020; Zhao and Zhang, 2020). However, very few researches on the influence of key match events on CSL matches currently exist. Therefore, the present study investigates the impact of the first goal on the match outcome (win, draw, and lose) and the goal difference (GD) in the Chinese Super League. Thus, three hypotheses are proposed: (a) scoring first increases a team's winning probability, (b) the first goal occurring later in a match increases a team's winning probability, and (c) the influence of the minute of the first goal and goal difference on the match outcome is also influenced by situational factors and the team's budget.

## MATERIALS AND METHODS

### Data Sources and Reliability

Technical performance data were collected from the website of Champion,<sup>1</sup> a leading Chinese sport data company. The validity and reliability of the Champdas Master Match Analysis System was verified by Gong et al. (2019), who demonstrated that Aiken's  $V$  averaged  $0.84 \pm 0.03$  and  $0.85 \pm 0.03$  for the validation of indicators. Moreover, high kappa values (Operator 1: 0.92 and 0.90; Operator 2: 0.91 and 0.88), high intraclass correlation coefficients (varied from 0.93 to 1.00), and low standardized typical errors (varied from 0.01 to 0.34) indicated a high level of intraoperator reliability. The kappa values for the interoperator reliability were 0.97 and 0.89.

Preseason budget data (transfer value) were collected *via* a website,<sup>2</sup> which was previously employed by Liu et al. (2019b). A positive correlation has been shown between the so-called "market values" determined by players' salaries and the teams' ranking (Bryson et al., 2013).

### Samples and Variables

From the 1,200 matches played in the 2014 to 2018 seasons in the CSL, a final sample of 1,116 matches in which at least one goal was scored were observed. The CSL is the highest-level national football league in China, and it includes 16 teams. These teams played in a balanced home and away schedule (15 home and 15 away matches for each team and season).

The dependent variables (**Table 1**) that were used were match outcome (MO) and the absolute goal difference (AGD). The independent variables were as follows: time (T), match

<sup>1</sup><http://data.champdas.com>

<sup>2</sup><http://transfermarkt.de>

**TABLE 1** | Description and categories of the variables.

Variables	Description
Match outcome (MO)	Outcome at the end of the match. It can end in a draw or in a victory for the local or visiting team
Absolute goal difference (AGD)	Goal difference at the end of the match
Time (T)	Time (in minutes) where the first goal is scored
Match location (ML)	Team playing like local or visiting
Team budget (TB)	Budget of the team scoring first (budget high or low)
Opponent budget (OB)	Budget of the opponent to the team scoring first (budget high or low)

**TABLE 2** | Description of the game results of the scoring first teams.

	Win	Draw	Lose	Total
<b>Match location</b>				
Home	464 (71.94%)	123 (19.97%)	58 (8.99%)	645
Away	276 (58.60%)	108 (22.93%)	87 (18.47%)	471
<b>Team Budget</b>				
Budget high	191 (76.10%)	42 (16.73%)	18 (7.17%)	251
Budget low	549 (63.47%)	189 (21.85%)	127 (14.68%)	865
Total	740 (66.31%)	231 (20.70%)	145 (12.99%)	1,116

location (ML), team budget (TB), and opponent budget (OB). In order to avoid estimation bias caused by endogenous variables, TB and OB were measured according to the rank of the preseason budget, considering that the preseason budget has a great association with the end-of-season ranking in football (Lago-Peñas and Sampaio, 2015; Liu et al., 2019b). All the observed criteria were applied to the team that scored first.

## Procedure and Statistical Analysis

First, after data were collected, they were ordered in specific spreadsheets, and information about match location, minutes when the goal was scored, and team budget was added. The teams were clustered into two groups in terms of the team's budget (high or low budget) using a k-means (Schwartz's Bayesian) technique. The higher group had a mean budget of 50.77 million dollars, while the lower group had a mean of 13.77 million dollars.

Second, the descriptive statistics (means and standard deviations) were calculated for the match outcome (MO) vs. match location (ML) and team budget (TB).

Third, in order to estimate the effects of the minute of the first goal on the match outcome, a multinomial logit regression model was adopted. The model can be specified as follows:

$$MO_{ij}(\text{win,draw,lose}) = \alpha_{0j} + \alpha_{1j}Time_{ij} + \alpha_{2j}ML_{ij} + \alpha_{3j}TB_{ij} + \alpha_{4j}OB_{ij} + \varepsilon_{ij} \quad (1)$$

where  $Y_{ij}$  is a random variable indicating the result of the  $i$ th match ( $j \in (1,2,3)$ , where  $1 = \text{win}$ ,  $2 = \text{lose}$ , and  $3 = \text{draw}$ ),  $\alpha_{0j}$  is the constant terms, and  $\varepsilon_{ij}$  is the disturbance terms that are independent and identically distributed following a Gumbel (type 1 extreme value) distribution. The equation

shows the influences of the independent variables related to scoring first on match outcome.

Additionally, a Poisson regression model was applied to further investigate the effects of scoring first on the absolute goal difference:

$$AGD_i = \beta_0 + \beta_{11}Time_i * D_1 + \beta_{12}Time_i * D_2 + \beta_{21}ML_i * D_1 + \beta_{22}ML_i * D_2 + \beta_{31}TB_i * D_1 + \beta_{32}TB_i * D_2 + \beta_{41}OB_i * D_1 + \beta_{42}OB_i * D_2 + \mu_i \quad (2)$$

where  $\beta_0$  is the constant term, and  $\mu_i$  is the error term for this regression model.  $D_1$  and  $D_2$  are dummy variables indicating whether a first scoring team eventually won the match.  $D_1 = 1$  if the first team that scored won the match, and  $D_2 = 1$  if the first team that scored did not win the match. Thus, the coefficients of the interaction terms in model (2) show the impacts of the variables related to scoring first on the absolute value of the goal difference, conditional on the match outcome for the first scoring team (i.e., either win or not win).

The abovementioned models were estimated using STATA for Windows (version 15.0, Stata Corp., Texas, United States), and the level of significance was set at  $p < 0.05$ .

## RESULTS

As **Table 2** shows, the majority of teams that scored first won the games (66.31%), followed by achieving a draw (20.70%) and losing (12.99%). Regarding the match location, the teams who scored first and played at home mostly won the matches (71.94%), followed by achieving draws (19.97%) and losing (8.99%). On the other hand, the teams who scored first and played away won the match in 58.60% of the cases, achieved a draw in 22.93% of the cases, and lost in 18.47% of the cases. Additionally, in terms of the team budget, when high-budget teams scored first, they won a higher percentage of their matches (76.10%) than low-budget teams (63.47%), whereas low-budget teams drew more matches (21.85%) than high-budget teams (16.75%).

The estimation results of the multinomial logit model are presented in **Table 3**. "Lose" was the omitted category (the basic outcome) with which the estimated coefficients are compared. The relative risk ratios are also included in this table to make the interpretation easier. All estimated parameters were statistically significant, indicating that the match location and team budget were the main factors.

In order to better understand the effects of the selected variables, the marginal effects of changing their values on the probability of observing the three match outcomes were, respectively, calculated. The values of the predicted probabilities indicated that the first scoring teams in the sample had the highest probability of winning (66.31%) and the lowest probability (12.99%) of losing matches. The probability of a draw was 20.70%. A statistically significant impact of the time of the first goal in the match-on-match outcome was found. The positive signs of its coefficients indicated that the later the goal scored, the higher the probabilities of winning or drawing. Moreover, the calculated average marginal effect demonstrated that postponing the first goal for 1 min increased the average

**TABLE 3** | Estimation results of the multinomial logit model.

Variables	$\beta$ Coefficient	Std. Err.	Relative-risk ratios	Marginal effects
<b>Lose</b>				Pr = 0.1299
Time of the first scoring				-0.0022***
Match location				-0.0952***
Team budget (Base outcome)				-0.0975***
Opponent budget				0.0647***
Constant				
<b>Win</b>				Pr = 0.6631
Time of the First Scoring	0.0232***	0.0049	1.0234	0.0028***
Match location	1.0083***	0.1910	2.7409	0.1342***
Team budget	1.0534***	0.2721	2.8674	0.1490***
Opponent budget	-0.6568***	0.2240	0.5185	-0.0752**
Constant	0.4254	0.1892	1.5303	
<b>Draw</b>				Pr = 0.2070
Time of the first scoring	0.0155***	0.0055	1.0156	-0.0006
Match location	0.5892***	0.2170	1.8026	-0.0390
Team budget	0.5501*	0.3072	1.7334	-0.0514*
Opponent budget	-0.4748*	0.2588	0.6220	0.0106
Constant	-0.1840	0.2154	0.8319	
N. of Obs.		1,116		
Log likelihood		-923.2150***		
LR chi2		81.15***		

The base outcome in this multinomial logit model is "Lose." "Pr" is the predicted probability calculated at sample mean value. \*Significance levels are indicated as  $p < 0.1$ . \*\*Significance levels are indicated as  $p < 0.05$ . \*\*\*Significance levels are indicated as  $p < 0.01$ .

probability of winning by 0.0028 while decreasing the probability of losing by 0.0022. Match location and the team budget of the first team that scored had a significant influence on game results. If the high budget team scores first playing at home, there were greater probabilities of winning or drawing than losing. Concretely, compared with away teams, home teams were approximately 13.42% more likely to win and approximately 9.52 and 3.90% less likely to lose and draw, respectively. High-budget teams were approximately 14.90% more likely to win and approximately 9.75 and 5.14% less likely to lose and draw, respectively, than low-budget teams.

The estimation results of the Poisson regression model are reported in **Table 4**. Most estimated coefficients in this model were statistically significant. To help interpret the parameter estimates of the Poisson regression, the marginal effects of all dependent variables for the probabilities of having a goal difference in observed matches were also listed.<sup>3</sup> The minute of the first goal had negative effects on the absolute value of the goal difference. For any team scoring first, the earlier the first goal

<sup>3</sup>The marginal effects of the explanatory variables for the probabilities of zero goal difference (in other words, the probabilities for the first scoring teams drawing) have the same magnitudes but opposite signs. Considering the main purpose of this paper, the following discussion will focus primarily on the probabilities of having a goal difference (i.e., the cases when the first scoring teams won or lost the match).

**TABLE 4** | Maximum likelihood estimation (MLE) results of the Poisson regression model.

Variables	Coef.	Std. Err.	Marginal effects for Pr (AGD $\geq$ 1)
Time for the first scoring * D1	-0.0044*	0.0012	-0.0012*
Time for the first scoring * D2	-0.0289*	0.0034	-0.0084*
Match location * D1	0.2287*	0.0558	0.0666*
Match location * D2	-0.7314*	0.1292	-0.2129*
Team budget * D1	0.2704*	0.0581	0.0787*
Team budget * D2	-0.5874	0.2045	-0.1710
Opponent budget * D1	-0.0352	0.0744	-0.0102
Opponent budget * D2	0.2382	0.1397	0.0693
Constant	0.5149*	0.0567	
N. of Obs.		1,116	
Log pseudolikelihood	-1,463.3526*		
Wald chi2	56.21***		

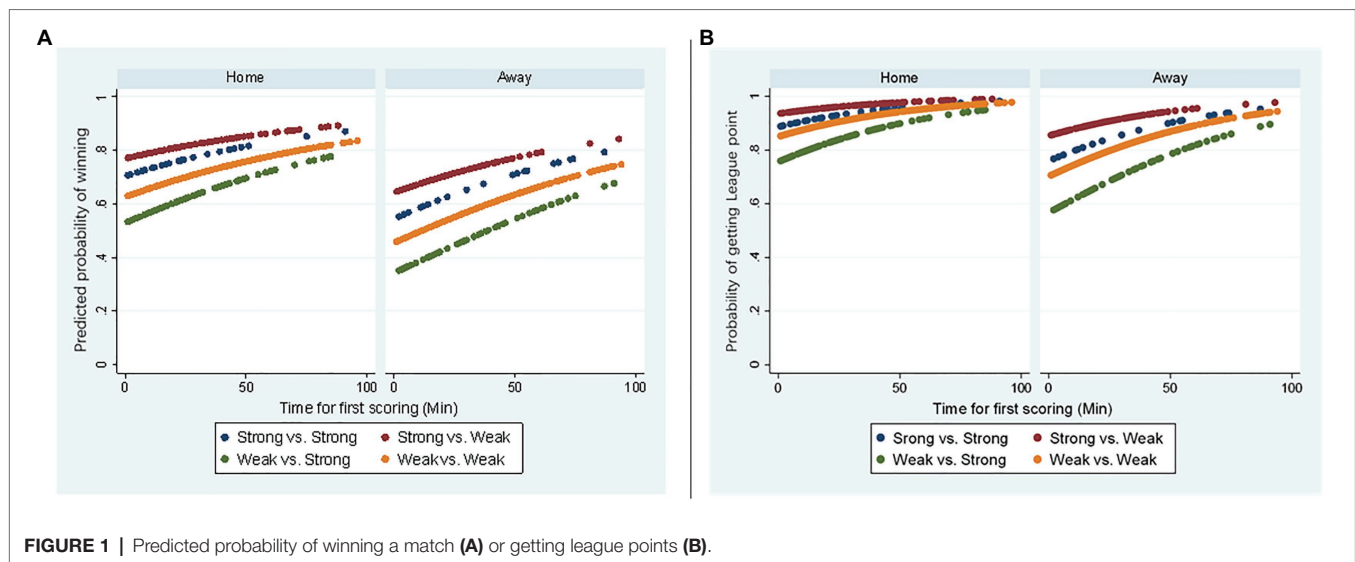
\*Significance levels are indicated as  $p < 0.01$ .

was scored, the larger the goal difference. Moreover, match location positively affected the goal difference. The opposite signs of the two interaction terms of match location and the dummies ( $D_1$  and  $D_2$ ) indicated that for the teams scoring first who won the match in the end, home teams had a higher likelihood of obtaining a final score with a larger goal difference than visiting teams. Furthermore, for those who did not win the match, home teams have a higher likelihood of obtaining a final score with a smaller goal difference than visiting teams. Besides, among the first scoring teams that won the match, teams with a high budget were more likely to obtain a final larger goal difference than those with a low budget.

The predicted probabilities of winning a match and obtaining league points for the first scoring teams were also graphed. At a given minute for the first goal, home teams had a higher probability of winning than away teams, while high-budget teams had a higher probability of winning than low-budget teams (**Figure 1A**). Furthermore, for any team, as the minute of the first goal moves forward to the end of the match, the predicted probability of winning increased accordingly. Furthermore, the team scoring first showed a high probability (>50%) of obtaining league points (**Figure 1B**).

## DISCUSSION

The aim of this study was to explore the effect of the minute of the first goal on the match outcome. The results found that (a) scoring first has a positive impact on the match outcome in the CSL. The teams scoring first have an 87.01% probability of being unbeaten and a probability of winning five times than of losing in the league. (b) The time of the first score in a match has a statistically significant impact on match outcomes. Delaying the first



**FIGURE 1** | Predicted probability of winning a match (A) or getting league points (B).

minute score for 1 min increases the average probability of winning for the first scoring team by 0.0028. (c) The impact of scoring first is also affected by the match location (ML) and team budget (TB). Thus, the higher the budget of the team that plays at home, the greater the probability of winning the match and achieving a higher goal difference if its first goal is scored at the beginning of the match. Therefore, the impact of scoring first is affected by contextual factors, such as the match location and team budget. The three hypotheses proposed are accepted.

The current research states that the influence of scoring the first goal in CSL was strong. Of all matches with scoring, 66.31% of the teams that scored first won their matches. Additionally, in all 1,200 matches, 80.92% of the teams that scored first won or drew at the end of the match. When playing at home, first scoring teams have an approximately 80% likelihood of obtaining league points and an over 50% win probability. When playing on the road, these results are over 50% and approximately 40%, respectively. This trend of the advantage of scoring first was consistent with previous results on both international and national events. In international tournaments (including FIFA World Cup), scoring first helped teams win over 70% of their matches (Leite, 2015; Vergonis et al., 2019). In domestic leagues, scoring first also increased the probability of winning (71.43–86.15%; Armatas, et al., 2009a,b; Molinuevo and Bermejo, 2012; Lago-Peñas et al., 2016). In the CSL, the 66.31% probability of winning for teams scoring first is less than that of listed tournaments, which is probably because at the event level, high-level tournaments such as the World Cup and UEFA Champions League have stronger teams. Scoring even one goal in a match between strong teams might be decisive, which possibly results in a high advantage of scoring first.

This advantage of scoring first was traditionally associated with tactical, technical, physical, and psychological factors. In football matches, those teams scoring first tend to change their playing style and strategy after scoring a goal (Bloomfield et al., 2005; Taylor et al., 2008). In order to prevent their opponent from scoring, they usually adopt a more conservative strategy (more defensive; Lago-Peñas and Gomez-Lopez, 2014).

Previous research also found that some of the performance indicators of teams scoring first are similar to successful teams (Szwarc, 2004). The reason could be that after the first goal is scored, players' psychological status changes. Compared with their opponents' pressure score after conceding a goal, a goal makes them feel relatively at ease in the following periods of the match and increases their confidence (Theis, 1992). Especially when playing at home, home teams enjoy home crowd support (Pratas et al., 2018), which negatively affects away teams (Wolfson et al., 2005). While scoring first improves a team's confidence, it also negatively impacts its opponent's psychology, consequently reducing its opponent's performance in competition (Leite and Barreira, 2014).

The regression model also reveals the following: (1) match outcome (win, draw, and loss) is influenced by the time of the first score (time), match location (ML), team budget (TB), and opponent budget (OB); and (2) the absolute goal difference (AGD) is influenced by the time of the first score (time), match location (ML), and team budget (TB). The later the first goal is made, the clearer the match outcome is determined. As stated in the *Introduction* section, in a low-scoring sport, the first goal is influential in determining the match result (Bloomfield et al., 2005; Lago-Peñas et al., 2016). In addition, a case study of Real Betis (a La Liga club) conceding goals reports that the last 15 min is decisive in the balance of goals conceded, and this influence increases if the team is losing (Méndez et al., 2020). Nevo and Ritov (2013, p.175) suggested that "if a team is leading or behind by a goal, their next goal is equally likely to be scored. While time passes with no goal being scored, the probability of a goal in the next move is [*sic*] increases." The current study found that it is partially true that the earlier the first goal of the team is, the larger the absolute goal difference (AGD). However, it depends on team strength and where the match is held (match location). If a team that has a high budget and plays at home scores first, it is very likely to continue scoring and win the match.

However, if a team that has a low budget and is playing in the road scores first at a very early time, this team may concede

more goals through the match. These findings are consistent with Pratas et al. (2018) who found that in the Portuguese Premier League, a first scoring home team would easily score a second goal according to the first scoring time. For an away team, having an advantage (specifically only one goal) will not naturally lead to a victory because the home team will not give up easily, especially when the home team is down by only one goal at halftime (Heuer and Rubner, 2012). Nevertheless, home field advantage is not the same for all teams (Clarke and Norman, 1995). In the Chinese Super League, strong and weak teams have different home field advantages. A long-term study (Zhou et al., 2020) found that strong opponents become continuously difficult to beat, which is probably due to their increasing budgets and the investment of rich CSL clubs. Thus, the home field advantage appears in matches with a balanced quality between teams, whereas unbalanced matches are favorable to the strong team regardless of the match location (Liu et al., 2019a). In the Spanish women's football professional league, the best teams are also highly effective regardless of the match location (Ibáñez et al., 2018). As a result, home field advantage is more likely to play a role in "plus value."

Applied in previous studies (Lago-Peñas and Gomez-Lopez, 2014; Lago-Peñas and Sampaio, 2015; Liu et al., 2019b), team budget was proven to have a major influence on teams' performance. Team budget indicates that rich clubs more easily sign star players and consequently organize a strong team. Good players are good in their physical (e.g., more high-intensity running, etc.; Mohr et al., 2003), tactical, and technical skills and their psychological ability against pressure (Lago-Peñas and Sampaio, 2015). These playing abilities help them adapt to the match situation and perform well and steadily on the pitch, even when facing a match status of losing. Moreover, players in successful teams perform more effectively and have higher quality of the tactical and technical skills rather than the number and quantity (Lago-Peñas et al., 2010; Yue et al., 2014; Liu et al., 2015). This could be the reason that strong teams can usually pull back when losing in an early stage in the match, and even they are able to get a big win in the end. However, weak teams need to change their strategies because they are not "strong" enough to keep their playing style under different situations. Furthermore, Mohr et al. (2003) reported that a reduced amount of high-intensity work toward the end of the match was linked to the match outcome. Strong teams have better physical conditioning, ensuring that they maintain high-quality technical and tactical actions (e.g., passing accuracy) and make fewer errors toward the end of match regardless of whether they are winning or losing. However, when weak teams are losing, their need for scoring strengthens their players' anxiety, resulting in increased offensive action but more errors (García-Rubio et al., 2015). This could be seized by strong teams, leading to conceding goal(s).

Despite these findings, the main limitation of this study is related to the small number of variables examined, the lack of information about technical-tactical performance indicators leading to goals, and substitutions or players' exclusion during games. Moreover, CSL clubs have increased their budgets and investment in recruiting talented players over these years, but this had not occurred for every club; and the long-term influence needs to be noticed. Last, team budget and match location

might have interacted to influence match results, which needs to be further considered. Future studies are suggested in order to merge all the contextual variables (especially the quality of the opposition) with the team budget and the main players signed by teams at the beginning of the season.

## CONCLUSIONS AND PRACTICAL APPLICATIONS

This study demonstrates that there is a strong advantage when a team scores first in the CSL. Even low-budget teams maintain a high probability of winning (63.47%). For home teams and high-budget teams, they have a very high probability of winning a game (home team 71.94% and high-budget team 76.10%) when they score first. The time of scoring the first goal has an impact on the match outcome, and scoring first at a later stage increases the win probability. The impact of scoring first is also influenced by contextual variables. When low-budget teams or visiting teams score first, they must focus more on their defense, or they may be more likely to lose. When high-budget teams or home teams score first, they tend to continue to score more. When high-budget teams or home teams concede a goal first, especially in the early phase of a match, they still have chance to equalize the score and even score many more goals to win. These results may help coaches develop proper strategies when facing different match statuses and different qualities of opposition in terms of team budget according to the match location.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <http://data.champdas.com>.

## AUTHOR CONTRIBUTIONS

TL and QC conceptualized the study. TL, AG-d-A, and QC contributed to the methodology and reviewed and edited the manuscript. QC contributed to the software, data collection, and visualization. TL wrote and prepared the original draft. HW and PH helped in improving this work. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry under a grant (KJZXCJ2015203).

## ACKNOWLEDGMENTS

The authors would like to thank Champion Co., Ltd. (Shanghai) whose data helped us to conduct this research.

## REFERENCES

- Almeida, C. H., Ferreira, A. P., and Volossovitch, A. (2014). Effects of match location, match status and quality of opposition on regaining possession in UEFA champions league. *J. Hum. Kinet.* 41, 203–214. doi: 10.2478/hukin-2014-0048
- Altarriba-Bartés, A., Calle, M. L., Susín, A., Gonçalves, B., Vives, M., Sampaio, J., et al. (2020). Analysis of the winning probability and the scoring actions in the American professional soccer championship. [Análisis de la probabilidad de ganar y de las acciones que conducen al gol en la liga americana de fútbol profesional]. *RICYDE. Rev. Int. cienc. deporte* 16, 67–84. doi: 10.5232/ricyde2020.05906
- Anderson, C., and Sally, D. (2013). *The Numbers Game: Why Everything You Know About Soccer Is Wrong*. London: Penguin Books.
- Armatas, V., Giannakos, A., Skoufas, D., Zangelidis, G., Papadopoulou, S., and Fragkos, N. (2009a). Goal scoring patterns in Greek top leveled soccer matches. *Citius Altius Fortius* 23, 46–51.
- Armatas, V., Yiannakos, A., Papadopoulou, S., and Skoufas, D. (2009b). Evaluation of goals scored in top ranking soccer matches: Greek “Super League” 2006–07. *Serb. J. Sports Sci.* 3, 39–43.
- Bar-Eli, M., Tenenbaum, G., and Geister, S. (2006). Consequences of players’ dismissal in professional soccer: a crisis-related analysis of group-size effects. *J. Sports Sci.* 24, 1083–1094. doi: 10.1080/02640410500432599
- Bloomfield, J., Polman, R., and O’Donoghue, P. (2005). Effects of score-line on team strategies in FA premier league soccer. *J. Sports Sci.* 23, 192–193.
- Bryson, A., Frick, B., and Simmons, R. (2013). The returns to scarce talent: footedness and player remuneration in European soccer. *J. Sports Econ.* 14, 606–628. doi: 10.1177/1527002511435118
- Clarke, S. R., and Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. *J. R. Stat. Soc. Ser. D* 44, 509–521.
- Ferreira, A. P. (2013). “From game momentum to criticality of game situations,” in *Routledge Handbook of Sports Performance Analysis*. eds. T. McGarry, P. O’Donoghue and J. Sampaio (New York, USA: Routledge), 288–300.
- Gai, Y., Volossovitch, A., Leicht, A. S., and Gómez, M.-Á. (2019). Technical and physical performances of Chinese Super League soccer players differ according to their playing status and position. *Int. J. Perform. Anal. Sport* 19, 878–892. doi: 10.1080/24748668.2019.1669356
- García-Rubio, J., Gómez, M. Á., Lago-Peñas, C., and Ibáñez, J. S. (2015). Effect of match venue, scoring first and quality of opposition on match outcome in the UEFA Champions League. *Int. J. Perform. Anal. Sport* 15, 527–539. doi: 10.1080/24748668.2015.11868811
- Garganta, J. (1997). *Modelação tática do jogo de futebol: estudo da organização da fase ofensiva em equipas de alto rendimento*. Porto: University of Porto Thesis in Portuguese.
- Gómez-Ruano, M., Carlos, L.-P., and Pollard, R. (2013). “Situational variables,” in *Routledge Handbook of Sports Performance Analysis. 1st Edn.* eds. T. McGarry, P. O’Donoghue, J. Sampaio and A. J. de Eira Sampaio (London, UK: Routledge).
- Gong, B., Cui, Y., Gai, Y., Yi, Q., and Gomez, M. A. (2019). The validity and reliability of live football match statistics from champdas master match analysis system. *Front. Psychol.* 10:1339. doi: 10.3389/fpsyg.2019.01339
- Han, B., Chen, Q., Lago-Penas, C., Wang, C., and Liu, T. (2020). The influence of the video assistant referee on the Chinese Super League. *Int. J. Sports Sci. Coach.* 15, 662–668. doi: 10.1177/1747954120938984
- Heuer, A., and Rubner, O. (2012). How does the past of a soccer match influence its future? Concepts and statistical analysis. *PLoS One* 7:e47678. doi: 10.1371/journal.pone.0047678
- Hucaljuk, J., and Rakipović, A. (2011). “Predicting football scores using machine learning techniques,” in *2011 Proceedings of the 34th International Convention MIPRO; May 23–27, 2011; (New York; IEEE)*, 1623–1627.
- Hughes, M., and Franks, I. M. (2004). *Notational Analysis of Sport: Systems for Better Coaching and Performance in Sport. 2nd Edn.* London: Routledge.
- Ibáñez, S. J., Pérez-Goye, J. A., Courel-Ibáñez, J., and García-Rubio, J. (2018). The impact of scoring first on match outcome in women’s professional football. *Int. J. Perform. Anal. Sport* 18, 318–326. doi: 10.1080/24748668.2018.1475197
- Jones, B. M. (2009). Scoring first and home advantage in the NHL. *Int. J. Perform. Anal. Sport* 9, 320–331. doi: 10.1080/24748668.2009.11868489
- Lago, C. (2009). The influence of match location, quality of opposition, and match status on possession strategies in professional association football. *J. Sports Sci.* 27, 1463–1469. doi: 10.1080/02640410903131681
- Lago-Peñas, C., and Dellal, A. (2010). Ball possession strategies in elite soccer according to the evolution of the match-score: the influence of situational variables. *J. Hum. Kinet.* 25, 93–100. doi: 10.2478/v10078-010-0036-z
- Lago-Peñas, C., and Gomez-Lopez, M. (2014). How important is it to score a goal? The influence of the scoreline on match performance in elite soccer. *Percept. Mot. Skills* 119, 774–784. doi: 10.2466/23.27.PMS.119c32z1
- Lago-Peñas, C., Gómez-Ruano, M., Megías-Navarro, D., and Pollard, R. (2016). Home advantage in football: examining the effect of scoring first on match outcome in the five major European leagues. *Int. J. Perform. Anal. Sport* 16, 411–421. doi: 10.1080/24748668.2016.11868897
- Lago-Peñas, C., Lago-Ballesteros, J., Dellal, A., and Gómez, M. (2010). Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league. *J. Sports Sci. Med.* 9:288.
- Lago-Peñas, C., and Sampaio, J. (2015). Just how important is a good season start? Overall team performance and financial budget of elite soccer clubs. *J. Sports Sci.* 33, 1214–1218. doi: 10.1080/02640414.2014.986184
- Leite, W. (2015). The influence of the first goal on the final result of the football match. *Balt. J. Sport, Health Sci.* 3, 29–35. doi: 10.33607/bjshs.v3i98.91
- Leite, W., and Barreira, D. (2014). Are the teams sports soccer, futsal and beach soccer similar. *Int. J. Sports Sci. Coach.* 2014:84. doi: 10.5923/s.sports.201401.11
- Liu, T. (2014). *Systematische spielbeobachtung im internationalen leistungsfußball*. doctoral dissertation. University of Bayreuth.
- Liu, T., Fournier-Viger, P., and Hohmann, A. (2018). “Using diagnostic analysis to discover offensive patterns in a football game,” in *Recent developments in data science and business analytics*. eds. M. Taviana and S. Patnaik (Cham, Switzerland: Springer). 381–386.
- Liu, T., García-De-Alcaraz, A., Zhang, L., and Zhang, Y. (2019a). Exploring home advantage and quality of opposition interactions in the Chinese Football Super League. *Int. J. Perform. Anal. Sport* 19, 289–301. doi: 10.1080/24748668.2019.1600907
- Liu, H., Gomez, M. A., Goncalves, B., and Sampaio, J. (2016). Technical performance and match-to-match variation in elite football teams. *J. Sports Sci.* 34, 509–518. doi: 10.1080/02640414.2015.1117121
- Liu, H., Gomez, M. A., Lago-Penas, C., and Sampaio, J. (2015). Match statistics related to winning in the group stage of 2014 Brazil FIFA World Cup. *J. Sports Sci.* 33, 1205–1213. doi: 10.1080/02640414.2015.1022578
- Liu, T., Ma, C., Zhang, D., Liu, J., and Bai, X. (2019b). Research on the correlation between team performance & season-end ranking and season-start club budget in Chinese Football Super League. *J. Capital Inst. Phys. Educ.* 31, 511–516. doi: 10.14036/j.cnki.cn11-4513.2019.06.007
- Mao, L., Peng, Z., Liu, H., and Gómez, M.-A. (2016). Identifying keys to win in the Chinese professional soccer league. *Int. J. Perform. Anal. Sport* 16, 935–947. doi: 10.1080/24748668.2016.11868940
- Martinez, F. D., and Gonzalez-Garcia, H. (2019). Effect of scoring first and match period in football world cup and Uefa Euro. *Eur. J. Hum. Mov.* 42, 123–137.
- Méndez, R. D. A., González-Jurado, J. A., and Otero-Saborido, F. M. (2020). Análisis observacional de goles recibidos por el Real Betis en La Liga: estudio de caso. Retos: nuevas tendencias en educación física, deporte y recreación. 355–362.
- Mohr, M., Krstrup, P., and Bangsbo, J. (2003). Match performance of high-standard soccer players with special reference to development of fatigue. *J. Sports Sci.* 21, 519–528. doi: 10.1080/0264041031000071182
- Molinuevo, J. S., and Bermejo, J. P. (2012). El efecto de marcar primero y la ventaja de jugar en casa en la liga de fútbol y en la liga de fútbol sala de España. *Rev. de Psicol. del Dep.* 21, 301–308.
- Nevo, D., and Ritov, Y. A. (2013). Around the goal: examining the effect of the first goal on the second goal in soccer using survival analysis methods. *J. Quant. Anal. Sport.* 9, 165–177. doi: 10.1515/jqas-2012-0004
- Pratas, J. M., Volossovitch, A., and Carita, A. I. (2016). The effect of performance indicators on the time the first goal is scored in football matches. *Int. J. Perform. Anal. Sport* 16, 347–354. doi: 10.1080/24748668.2016.11868891
- Pratas, J. M., Volossovitch, A., and Carita, A. I. (2018). Analysis of scoring sequences in matches of the Portuguese premier league. *J. Hum. Kinet.* 64, 255–263. doi: 10.1515/hukin-2017-0199

- Ruiz-Ruiz, C., Fradua, L., Fernández-García, Á., and Zubillaga, A. (2013). Analysis of entries into the penalty area as a performance indicator in soccer. *Eur. J. Sport Sci.* 13, 241–248. doi: 10.1080/17461391.2011.606834
- Sasaki, Y., Nevill, A., and Reilly, T. (1999). Home advantage: a case study of Ipswich town football club during the 1996–97 season. *J. Sports Sci.* 17: 831.
- Szwarc, A. (2004). Effectiveness of Brazilian and German teams and the teams defeated by them during the 17th Fifa World Cup. *Kinesiol. Int. J. Fundam. Appl. Kinesiol.* 36, 83–89.
- Taylor, B. J., Mellalieu, D. S., James, N., and Barter, P. (2010). Situation variable effects and tactical performance in professional association football. *Int. J. Perform. Anal. Sport* 10, 255–269. doi: 10.1080/24748668.2010.11868520
- Taylor, J. B., Mellalieu, S. D., James, N., and Shearer, D. A. (2008). The influence of match location, quality of opposition, and match status on technical performance in professional association football. *J. Sports Sci.* 26, 885–895. doi: 10.1080/02640410701836887
- Tenga, A. (2013). “First goal and home advantage at different levels of play in professional soccer,” in *Performance Analysis of Sport IX*. eds. D. M. Peters and P. O’Donoghue (New York: Routledge), 71–75.
- Theis, R. (1992). “Analyse von Torerfolgen im mittleren und unteren Amateurbereich im Fußball,” in *Analyse und Beobachtung in Training und Wettkampf*. eds. W. Kuhn and W. Schmidt (Sankt Augustin: Academia Verlag), 173–174.
- Tucker, W., Mellalieu, D. S., James, N., and Taylor, B. J. (2005). Game location effects in professional soccer: a case study. *Int. J. Perform. Anal. Sport* 5, 23–35. doi: 10.1080/24748668.2005.11868325
- Vergonis, A., Michailidis, Y., Mikikis, D., Semaltianou, E., Mavrommatis, G., Christoulas, K., et al. (2019). Technical and tactical analysis of goal scoring patterns in the 2018 Fifa World Cup in Russia. *FU Phys. Ed. Sport* 17, 181–193. doi: 10.22190/fupes190612019v
- Wolfson, S., Wakelin, D., and Lewis, M. (2005). Football supporters’ perceptions of their role in the home advantage. *J. Sports Sci.* 23, 365–374. doi: 10.1080/02640410400021567
- Yang, G., Leicht, A. S., Lago, C., and Gómez, M.-Á. (2018). Key team physical and technical performance indicators indicative of team quality in the soccer Chinese super league. *Res. Sports Med.* 26, 158–167. doi: 10.1080/15438627.2018.1431539
- Yiannakos, A., and Armatas, V. (2006). Evaluation of the goal scoring patterns in European championship in Portugal 2004. *Int. J. Perform. Anal. Sport* 6, 178–188. doi: 10.1080/24748668.2006.11868366
- Yu, Q., Gai, Y., Gong, B., Gómez, M.-Á., and Cui, Y. (2020). Using passing network measures to determine the performance difference between foreign and domestic outfielder players in Chinese Football Super League. *Int. J. Sports Sci. Coach.* 15, 398–404. doi: 10.1177/1747954120905726
- Yue, Z., Broich, H., and Mester, J. (2014). Statistical analysis for the soccer matches of the first Bundesliga. *Int. J. Sports Sci. Coach.* 9, 553–560. doi: 10.1260/1747-9541.9.3.553
- Zhao, Y., and Zhang, H. (2020). Eigenvalues make the difference – A network analysis of the Chinese Super League. *Int. J. Sports Sci. Coach.* 15, 184–194. doi: 10.1177/1747954120908822
- Zhou, C., Gómez, M.-Á., and Lorenzo, A. (2020). The evolution of physical and technical performance parameters in the Chinese Soccer Super League. *Biol. Sport* 37:139. doi: 10.5114/biolSport.2020.93039
- Zhou, C., Hopkins, W. G., Mao, W., Calvo, A. L., and Liu, H. (2019). Match performance of soccer teams in the Chinese Super League-effects of situational and environmental factors. *Int. J. Environ. Res. Public Health* 16:4238. doi: 10.3390/ijerph16214238
- Zhou, C., Zhang, S., Lorenzo Calvo, A., and Cui, Y. (2018). Chinese soccer association super league, 2012–2017: key performance indicators in balance games. *Int. J. Perform. Anal. Sport* 18, 645–656. doi: 10.1080/24748668.2018.1509254

**Conflict of Interest:** PH was employed by company Microsoft, Beijing.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Liu, García-de-Alcaraz, Wang, Hu and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.