Journal of Physical Education and Sport ® (JPES), Vol. 25 (issue 2), Art 38, pp. 338 - 346, February 2025
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Forecasting of football match results using econometric models

SEBASTIAN MAJEWSKI[*1], AGNIESZKA MAJEWSKA[2], ABDULLAH YİĞİT GÜNGÖR[3]
[1,2]Department of Sustainable Finance and Capital Markets, University of Szczecin, POLAND
[3]Department of Business Administration, Atatürk University, TÜRKİYE

Published online: February 28, 2025
Accepted for publication: February 15, 2025
DOI:10.7752/jpes.2025.02038

Abstract:
The issue of foresight in human activity plays a non-trivial role, particularly when financial considerations are involved. The ability to anticipate future events can provide an investor with a sense of security regarding the investment or market advantage over other market participants. There appears to be no such model that can be successfully used for forecasting and planning. Using the terminology of financial markets, the bookmaking market would be classified as an efficient market if such a model existed. Inefficiency leads to inequality and therefore an informational advantage for some market participants over others, which should result in the possibility to realize supernormal gains. Even from the perspective of market organizers, inefficiency that induces investors to gamble or increases the bookmaker's revenue seems more advantageous. The primary aim of this study is to test the feasibility of building a robust econometric model against which predictions of match results can be made. Thus, an indirect verification of the intermediate efficient market hypothesis for the sports betting market will be carried out. The study will be conducted concurrently for two autonomous football leagues: the Polish Ekstraklasa and the Türkiye 1 Super Futbol League during the years of 2021 and 2022. A series of panel regression models were utilized to test the hypothesis. The results of the analyses showed that it was possible to estimate the probability of winning (bookmakers' expectations).
Key Words*: Football, betting market, match results, efficient market, econometrics, travel effect*

Introduction

The global sports betting market size is $83.6 billion in 2022 as one of the fastest-growing markets in the world and it is expected to reach $182.12 billion by 2030 (www.grandviewresearch.com). Therefore, this market engages the attention of scholars and increases its popularity. Sports betting refers to putting wagers on outcomes of particular sports events that are associated with a specific profit predefined by a bookmaker. If the outcome is presumed right by the bettor, the wager is gotten back with its profit. Otherwise, the bookmaker retains the wager without any reimbursement (Hubáček et al., 2019).

Football has the highest share in the sports betting industry in the world and fixed odds betting is one of the most widely preferred styles in football betting markets. This type of betting entails the bookmaker's determination of the odds for the home team to win, the away team to win, and the draw. For this purpose, the fixed odds are identified by the bookmaker according to many indicators nearly a week prior. Hence, the bettor is aware of the potential profit when the wager is made (Dixon & Coles, 1997).

The analyses found in the field of sports economics, particularly on issues related to professional football, can be divided into several trends:

- Player valuation - the actual valuation of intangible assets in the form of player card rights (Carmichael et al., 1999; Dobson et al., 2000; Lucifora & Simmons, 2003; Turnau et al., 2005; Frick, 2011; Trequattrini et al., 2012; Wicker et al., 2013; Buriamo et al., 2015; Kesenne, 2015; Majewski, 2016; Muller et al., 2017; Güngör, 2024);
- The analysis of the links between stock returns and sports performance - there are several strands here including sports stocks, stock indices, or shares of football club sponsoring companies (Berument et al., 2009; Majewski & Majewska, 2018; Hagen & Cunha, 2019; Škrinjarić, & Barišić, 2019; Galloppo & Boido, 2020; Soti et al., 2020);
- Modelling and forecasting of football match results (Crowder et al., 2002; Goddard, 2005; Boshnakov et al., 2017; Holmes & McHale, 2023);
- Media, popularity, and other issues (Boyle, 2009; McCarthy et al., 2014; Szymanski & Weimar, 2019; Gövdeli & Güngör, 2022; Malagon-Selma et al., 2023).

Courneya and Carron (1992) have established a framework for game location research. This framework includes a process that shows how the match results are influenced by game location (home and away), game location factors (crowd, familiarity, travel, rules), critical psychological states (competitors, coaches, officials), critical behavioral states (competitors, coaches, officials). According to the framework, one of the important

Corresponding Author: SEBASTIAN MAJEWSKI, E-mail: sebastian.majewski@usz.edu.pl

factors is the travel effect which is experienced by away teams. It can induce tiredness, jet lag, or fatigue in away teams, hence influencing the match results.

Fama (1970) suggested efficient market theory and defined it as the reflection of all information on the prices. In the simplest definition of market efficiency, the market is deemed efficient if security prices quoted on this market fully reflect all available information. The efficient market can be classified as weak form, semi-strong form, and strong form. The weak form of market efficiency postulates an investor cannot have a return above average by using previous price movements. Semi-strong market efficiency postulates an investor cannot have a return above average by using previous price movements as well as publicly available information. Strong market efficiency postulates that nobody can have a return above average. According to theory, all information (publicly available and unavailable) is disseminated to all investors rapidly. Thus, the information reflects on pricing expeditiously (Fama, 1991).

This rather specific phenomenon is betting market efficiency. The requirement that expected returns be equal across betting prospects is the definition of efficiency that is the most stringent. In order to guarantee market efficiency, several concepts must be met. Along with others, they are (Sauer, 1998):

- Constant returns on various wagers;
- The lack of a profit opportunity and effective point spreads;
- Equilibrium pricing functions.

The early academic works on the efficiency of the betting market were mainly concerned with betting on horse racing. Snyder argues that if horserace betting markets are weakly efficient, 'then the expected rate of return for all types of bets would be identical' (Snyder, 1978). In the work of Williams (1999), it could be seen that the existence of semi-strong efficiency in betting markets would imply, therefore, that the expected returns to any bet, or type of bet, placed regarding different events, with identical probabilities of success, should themselves be identical (subject to equivalent costs and risks). The same applies with respect to strong efficiency when assessed with respect to all information.

Typically, inefficiencies in markets, such as equity markets, are associated with market anomalies that negate the possibility of such a market. The bookmaking market is no different, except that there are significantly fewer such market occurrences. One of the most widely described in the literature is the effect called favorite-longshot bias. The favorite-longshot bias, which was first observed in the horse race betting market, states that horses with lower odds (higher winning probability) tend to win more often than predicted by their odds, while horses with higher odds (lower winning probability) tend to win less frequently than predicted by their odds (Chung & Hwang, 2010). This implies that favorites are underbet while longshots are overbet in a parimutuel market structure (Thaler & Ziemba, 1988), a betting system in which all bets of a given type are placed together in a pool.

An understanding of the phenomena and processes that enable the sports betting market to function is required to understand how the probabilities of realization of events in the sports betting market are calculated. The market of bookmakers is divided into four categories (Jedraszka & Zaton, 2011):

- Ground - providing bets through their physical locations;
- Internet - providing bets online;
- Mixed - providing bets both physically and online;
- The segment of sports betting exchanges - the market for trading in probabilities, where a gambler has the option to both buy and sell bets on sporting events.

The probability of subjective forecasts, commonly referred to as betting odds, is the one factor that all of these parts have in common (Ayton, 1997). Such an interpretation of probability contradicts the assumptions of probability theory.

Three fundamental criteria are used to determine betting odds:

- Decimal - betting odds are decimal fractions greater than 1 - in this case, the probabilities are the inverse of betting odds;
- Fractional - betting odds are common fractions;
- Moneyline - betting odds are integral numbers with signs '+' or '-'.

Europe, with the exception of Great Britain and Ireland, uses the decimal system to determine betting odds. It only takes multiplying the betting odds by the amount to gain access to information about betting odds in this more open manner.

All bookmakers generate money through profit margins, and the amount they make is largely influenced by the following (Jedraszka & Zaton, 2011):

- The status of a sporting event;
- The type of business conducted;
- The country where taxes are levied;
- The margin policy.

Due to the data's accessibility and universality – betting odds published online allow for access to the information even if an investor does not gamble – betting odds acquired from bookmakers' websites were used in this research. The assumption behind this decision is that the same group of investors will participate in both of these marketplaces. These investors are traders, or participants, who view the market as one of many potential sources of income. Traders in both markets use the same methodical and professional approach to investing – seeking a competitive advantage in the deluge of information to outperform other market participants. Unfortunately, this may also be a factor in the transfer of risk from one market to another since information about event probabilities may have an impact on stock quotes for the football firm. These investigations, which were carried out in Europe, have demonstrated that there are a few statistically significant relationships between sporting outcomes, betting odds (anticipated probability), and the rates of return of football companies listed on European exchanges (Majewski, 2014a).

Material & methods

Previous work using quantitative methods to model data from sport events can be divided from the point of view of the tools used:

- Linear and nonlinear regression;
- Panel regression;
- Structural models;
- Autoregressive methods (ARCH, GARCH, GARCH-M, EGARCH, TARCH, etc.);
- Statistical tests;
- Granger's causality and co-integration tests.

Table 1. Brief Presentation of Some Examples of Quantitative Tools and Methods Run on Sports Data During the Last 30 Years

Goal of the research	Type of analysis	Authors
Valuation of stock exchange assets (prices, rates of return, values of index)	Linear and nonlinear regression	Stadtmann (2003); Bell et al. (2010); Aglietta et al. (2010); Demir & Daniş (2011); Saraç & Zeren (2013)
	GMM	Ashton et al. (2003)
	ARCH-GARCH	Douque & Ferreira (2005); Berument et al. (2006)
	EGARCH	Benkraiem et al. (2010); Berument & Ceylan (2012)
	Statistical tests	Bell et al. (2012)
	Granger's co-integration and causality tests	Leitão et al. (2012)
	Event analysis (OLS regression)	Chen et al. (2019)
	Panel regression	Abbas (2022)
	Dynamic multivariate model	Koopman & Lit (1019)
Game results	GARCH	Edmans et al. (2007)
	Linear and nonlinear regression	Klein et al. (2009); Goddard (2005)
	Panel regression	Baur & McKeating (2009);
	Probit regression	Goddart & Asimakopoulos (2004)
	Random forest estimation	Goller et al. (2021)
	Ordered logit regression	Arnzen & Hvattum (2021)
Valuation of players	Linear and nonlinear regression	Carmichael et al. (1999); Gerrard & Dobson (2000); Lucifora & Simmons (2003); Majewski (2014b); Kesenne (2015); Wicker et al. (2013); Majewski (2016)
	Structural model	Samagio et al. (2009)
	Trinomial tree	Turnau et al. (2005)
	DCF model	Trequattrini et al. (2012)

We decided to use panel regression models in this research because of the character of the data we collected. Baltagi uses the term "panel data" as a pool of observations on cross-sectional entities like i.e. households, countries, firms, etc. over several periods (Baltagi, 2005).

The classic form of the equation for panel data is always characterized by double indexing. Therefore, it can be written that the model using panel regression takes the following form:

$$y_{it} = \alpha + \beta x_{it} + u_{it}$$

where:

i – represents cross-sectional dimension (i.e. football clubs);
t – represents time series dimension (i.e. match week, round, etc.)
α – scalar representing estimator of intercept, β – vector of estimators for K explanatory variables;
x_{it} is the matrix of ith observations of independent variables from time t on K explanatory variables.

One-way error component model is used in most of panel data analysis:

$$u_{it} = \mu_i + v_{it}$$

where:
μ_i – denotes undetectable, individualized effects that are constant over time;
v_{it} – stands for the remaining unseen effects that change across people and across time.
There is also a two-way error component method used could be represented by an equation:

$$u_{it} = \mu_i + \lambda_t + v_{it}$$

Where λ_t is assigned to unobservable time-specific effects.
There are a variety of different types of models for panel data. They could be grouped into four classes (Green, 2003):

- pooled regression;
- fixed effects;
- random effects;
- random parameters.

The fixed effects model – Last Square Dummy Variable model (LSDV) was proposed by Green in 1993. In case when there is unobservable variable Z correlated with the set of independent variables X the estimators of regression parameters are biased and inconsistent, so the proposed estimation of the model is described by the following equation (Green, 2003):

$$y_{it} = \beta x_{it} + \alpha + u_{it}$$

where $\alpha = z_i\alpha$ and represents all observable and unobservable effects by the estimated mean. The term 'fixed' is specific to this group of regression models and it does not mean that z_i is stochastic but only that there exists a correlation between z_i and x_{it}.

The random effects model is possible to be estimated using a two-step Feasible General Least Squares (FGLS) method. Various combinations of the residual variances from the linear model with no effects, the group means regression, and the dummy variables result in a wide range of reliable estimators of the variance components. If it can be assumed that the included variables have no correlation with the unobserved individual heterogeneity, the model can be written as follows (Baltagi, 2005):

$$y_{it} = \beta x_{it} + \alpha + u_{it} + \varepsilon_{it}$$

Theoretically, such kind of model represents linear regression with a compound disturbance that may be consistently but inefficiently, estimated by the least squares method similar to the fixed effects model. In this approach, random effects are represented by two variables u_i and ε_{it}. With the exception of the fact that there is only one draw for each group and that draw enters the regression in the same way every period, the parameter u_i is a group-specific random element, similar to ε_{it}. The key distinction between fixed and random effects is not based on whether or not they are stochastic, but rather whether the unobserved individual effect has components that are associated with the model's regressors.

We have used only two simple approaches to conduct panel data analysis - fixed effects and random effects models. The sample of the study consists of the Polish First Tier Football League (Ekstraklasa) and the Turkish First Tier Football League (Süper Lig) for the 2021/2022 season. Ektraklasa had 18 and Super League had 20 teams for the particular domestic season. 306 matches were examined for Ekstraklasa and 380 matches were examined for Süper Lig. Consequently, the study's total sample includes 38 teams and 686 matches. Table 2 displays the information about the research sample.

Table 2. Sample of The Study

League	Season	Number of Teams	Number of Matches
Polish Ekstraklasa	2021/2022	18	306
Turkish Super League	2021/2022	20	380

The variables of the study are "travel distance", "value difference", "win", "loss", "points", and "betting odds". To measure the variable of travel distance, the stadium locations of the football clubs were firstly determined by information from Transfermarkt (www.transfermarkt.com). Then, the European Union distance calculator was used to calculate the travel distance between home and away teams (www.erasmus-plus.ec.europa.eu). Similarly, related information about the variables of win, loss, and points were obtained from Transfermarkt (www.transfermarkt.com). Lastly, pertinent information regarding the variable of betting odds was obtained from Oddsportal (www.oddsportal.com).
As a result of the econometric analyses, we obtained a series of panel regression models, the best of which will be presented in the tables below.

Table 3. Results of Estimation of Fixed Effects Model for Polish Ekstraklasa for The Probability of Winning (Bookmakers' Expectation of Winning)

Variable	Coefficient	Standard error	t	p-value	
const	0.343292	0.005629	60.98	<0.0001	***
TravDis	0.000318	2.62958e-05	12.09	<0.0001	***

ValDiff	0.000487	0.000281	1.734	0.0835	*
Estimation parameters:					
Mean Y	0.386897	St. deviation Y	0.154756		
The sum of residuals' squares	6.768985	Standard error	0.106930		
LSDV R-square	0.537421	Within R-square	0.201215		
LSDV F(19, 592)	36.19899	P-value for test F	4,00e-86		
Likelihood logarithm	509.9503	AIC	−979.9005		
SIC	−891.5659	HQIC	−945.5442		
Autocorrelation	0.011542	DW	1.913387		

The estimation of the fixed effects model was run using 612 observations for 18 cross-sectional data (sports clubs) and the time series length spanning 34 match weeks. The joint test on named regressors (p-value = P(F(2, 592) > 74,563) = 1,31579e^{-29}) and free expression variation test ingroups (p-value = P(F(17, 592) > 31,6106) = 9,84751e^{-72}) indicate a lack of reason to reject the null hypothesis. The LSDV model was fitted to the real data in 53,7% and both independent variables: travel distance (TravDis) and difference in market values (ValDiff) are positively correlated with the probability of winning and are statistically significant.
Both the model with the probability of winning and the model with the probability of losing show statistically significant parameters, but a slightly better fit is shown by the model with the probability of winning (see Table 4).

Table 4. Results of Estimation of Between Effects Model for Polish Ekstraklasa for The Number of Points Scored by The Team in The Single Match.

	Coefficient	Standard error	t	p-value	
const	0.088177	0.200584	0.4396	0.6661	
Winpro	3.298560	0.501847	6.573	<0.0001	***
Estimation parameters:					
Mean of dep. variable	1.364379	St dev. if dependent var.	0.398595		
Sum of rests squares	0.729954	Standard error	0.213593		
R-square	0.729740	Adjusted R-square	0.712849		
F(1, 16)	43.20227	p-value for F test	6.42e-06		
Likelihood logarithm	3.305420	AIC	−2.610840		

The model indicates the strong relationship between bookmakers' expectations about the winning of the match and the number of points scored by the team. The R-square describes this relationship in nearly 73%.
The best approximations for Turkish Leagues were obtained for two random effects models describing the probability of winning and points scored by the team in the match. The results are shown in Tables 5 and 6.

Table 5. Results of Estimation of Random Effects Model for Turkish League for The Probability of Winning (Bookmakers' Expectation of Winning)

	Coefficient	Standard error	z	p-value	
const	0.339113	0.015937	21.28	<0.0001	***
ValDiff	0.000135	6.78351e^{-05}	1.997	0.0458	**
TravDis	0,001190	7.57035e^{-05}	15.73	<0.0001	***
Estimation parameters:					
Mean Y	0.387317	Standard deviation Y	0.144242		
Sum of residuals' squares	13.31819	Standard error	0.132553		
Likelihood logarithm	458.3979	AIC	−910.7959		
SIC	−896.8959	HQIC	−905.4432		
Autocorrelation	−0.087953	DW	2.066024		

Estimation random effects (GLS) used 760 observations for 20 cross-sectional data units (number of teams in the Turkish League) and the time series length spanning 38 match weeks.

The quasi-demeaning methods give the following results: 'between' variance = 0.0045425, 'within' variance = 0.0109378 and theta = 0.75589.

The joint test on named regressors does not allow for rejection of the null hypothesis (Chi-square (2) = 250.308 with p-value = $4.43e^{-55}$). The statistic of the Breusch-Pagan test is 1991.32, indicating the no random effects hypothesis is rejected at the 5% level. The p-value for the Hausman test is less than 5%, suggesting that the random effects estimator is inconsistent. The conclusion from these tests is that even though there is evidence of random effects (LM rejects), the random effects are not independent of the regressors; the FGLS estimator will be inconsistent.

Similarly to models of Polish Ekstraklasa model uses travel distance (TravDis) and difference in market values (ValDiff) as independent variables, which are positively correlated with the probability of winning and are statistically significant.

Table 6. Results of Estimation of Random Effects Model for Turkish League for The Number of Points Scored by The Team in The Single Match.

	Coefficient	Standard error	z	p-value	
const	−0.09488	0.137453	−0.6903	0.4900	
Winpro	3.80522	0.322864	11.79	<0.0001	***
Estimation parameters:					
Mean Y	1.378947	Standard deviation Y	1.324176		
Sum of residuals' squares	1102.111	Standard error	1.205013		
Likelihood logarithm	−1219.626	AIC	2443.252		
SIC	2452.518	HQIC	2446.820		
Autocorrelation	0.045570	DW	1.872412		

Estimation random effects (GLS) used 760 observations for 20 cross-section data units included and the time series length was 38 match-weeks.

The quasi-demeaning methods give the following results: 'between' variance = 0.0275434, 'within' variance = 1.43101 and theta = 0.240023.

The joint test on named regressors does not allow for rejection of the null hypothesis (Chi-square (1) = 138.906 with p-value = $4.61709e^{-32}$). The statistic of the Breusch-Pagan test is 2.9694 indicates the no random effects hypothesis is accepted at the 5% level (p-value= 0.0848531). The p-value for the Hausman test is more than 5% which suggests that the random effects estimator is consistent (p-value = 0.992362).

Discussion

Some recent studies in the literature support the impact of the travel effect on football match results (e.g. Beckmann, 2022; Duran et al., 2021). These studies include a great number of matches, various football leagues, and different time durations. For instance, Beckmann (2022) examined the effect of travel on match results in football from 1964 to 2020 within the Bundesliga (German 1st division). The data includes 17,376 games and the findings support the travel effect on the outcome of the football matches which is consistent with our study. On the other hand, he also indicated that even though the effect was significant, the degree of the impact of travel distance had been decreasing gradually. In addition, the effect saturated around 100 km distance which was congruent with previous studies conducted in Italy, Türkiye, and England. Duran and his colleagues (2021) remarked on the unfairness of the travel effect on Argentinean youth leagues which consist of major divisions (Under 20, 18, and 17) and minor divisions (Under 16, 15, and 14). Both divisions had been played in the single-round format. However, the schedule of the minor divisions was the reverse matching of the major ones. This format caused a significant disadvantage in terms of the travel effect which is implicitly congruent with our study. Based on this problem, they proposed some models developed via mathematical programming for fair scheduling. Hence, the travel effect would be distributed more equitably.

Contrary to supportive studies, Gilbert and his colleagues (2020) hypothesized that “there will be a negative relationship between the number of hours travelled before a match and the match outcomes of visiting teams”. The hypothesis was tested in Major League Soccer (MLS) for the seasons of 2015-2016 and 2016-2017 which consisted of 1,054 games in total. An ordinal regression model was used for analysis in the study. However, the result did not support the hypothesis which is contentious to our study. In this case, the reason might be the discrepancy between the areas of the samples. The USA has an approximately 10 million km2 area while Poland and Türkiye cover 1 million km2 in total. Therefore, MLS is implementing strategic scheduling to mitigate the travel effect. For this purpose, away matches in similar regions are closely paired so the teams have opportunities to camp and travel shorter distances. On the other hand, they play two or three matches at home consecutively. This strategic scheduling might be useful for Polish Ekstraklasa and Turkish Super League.

Some studies in the recent literature support the effect of team value on football match results (e.g. Metelski et al., 2022; Warias and Block, 2021; Gerhards and Mutz, 2016). These studies pertain to both national/club team squads and long/short periods. In this context, Metelski and his colleagues (2022) explored the relationship between match results and the squad value of national teams that participated in the 2020 European Football Championship (EURO 2020). 24 national teams that competed in the tournament comprised the sample of the study. Pearson's correlation test denoted that the value of the squad had a substantial effect on match results. This finding is also coherent with our study. Warias and Block (2021) probed the effect of team value on football match results. They collected the data from top five European Leagues for the years 2014-2020. 3,067 matches were analysed and the regression findings showed that the value difference between teams was a decisive factor on match results. Lastly, Gerhards and Mutz (2016) studied some determinants of football match results. The data were collected from 12 European football leagues for the years 2011-2016. Regression analysis showed that team value is the most influential predictor by far. The findings also indicated that the impact of the team value is greater when the league has lower financial equity.

In summary, our study includes both travel effect (distance) and team value as the independent variables in the same model. In addition, the time period of our study (2021-2022) is more up-to-date, allowing us to test this model after the Covid-19 lockdown. The sample of the study (Polish Ekstraklasa and Turkish Super League) also increases the originality, since previous studies primarily focused on the top 5 European football leagues. Lastly, our study tested the hypothesis via the panel regression analysis which considers the time effect, further differentiating our study from previous ones.

Conclusions

Regardless of which league the econometric analyses were performed for using models for panel data, it can be unequivocally stated that:

- It is possible to estimate the probability of winning (bookmaker's expectations) by taking into account two basic variables: the difference in the value of the clubs playing the match (ValDiff) and the distance the club has to travel to play the match (TravDis). The parameters standing next to both variables for both the Polish and Turkish leagues had a positive sign which means that both variables have a positive effect on the value of bookmakers' expectations.
- In the case of the estimation concerning the number of points obtained by the club in a single match, only the probabilities of winning (or losing, although the latter were less well-fitted) were significant variables. Again, the sign of the parameter was positive, meaning that the higher the probability of winning (the bookmaker's expectation of winning) the more points the club was likely to score in a match.

Thus, treating these models as two equations solving the problem of predicting the outcome (win, draw, or lose), it can be said that there is a strong statistical relationship between the studied characteristics, and by using simple parameters, it is possible to make predictions of match results to a certain extent (the coefficient of determination R-square was greater than 50%).

Based on the conclusion of our study, we have suggestions for future studies as follows:

- New algorithms could be developed which are supported by artificial intelligence to schedule competition matches more fairly in terms of the travel effect.
- The effect of camping for away matches could be investigated regarding whether it mitigates the travel effect or not.
- The method of travel (highway, seaway, or airway) could be compared with each other.
- The saturation degree of the team value difference could be studied.
- First XI value could be tested instead of the squad value.
- Other variables could be added to the research model such as the crowd effect and the sample could be differentiated in terms of leagues, tournaments, and even gender.

Conflicts of interest - The authors have no conflict of interest to declare.

References

Arntzen, H., & Hvattum, L. M. (2021). Predicting match outcomes in association football using team ratings and player ratings. *Statistical Modelling*, *21*(5), 449-470.

Ayton, P. (1997). How to be incoherent and seductive: Bookmakers' odds and support theory. *Organizational Behavior and Human Decision Processes*, *72*(1), 99-115.

Baltagi, B. H. (2005). *Econometric analysis of panel data*, John Wiley & Sons.

Beckmann, N. (2022). Statistical influence of traveling distance on home advantage over 57 years in the men's German first soccer division. *German Journal of Exercise and Sport Research, 52*(4), 657-665.

Berument, M. H., Ceylan, N. B., & Ogut-Eker, G. (2009). Soccer, stock returns and fanaticism: Evidence from Turkey. *The Social Science Journal, 46*(3), 594-600.

Boshnakov, G., Kharrat, T., & McHale, I. G. (2017). A bivariate Weibull count model for forecasting association football scores. *International Journal of Forecasting*, *33*(2), 458-466.

Boyle, R. (2009). *Power play: Sport, the media and popular culture*. Edinburgh University Press.

Buriamo B., Frick B., Hickfang M., & Simmons R. (2015). The economics of long-term contracts in the footballers' labour market. *Scottish Journal of Political Economy*, *62*(1), 8-24.

Carmichael, F., Forrest, D., & Simmons, R. (1999). The labor market in association football: Who gets transferred and for how much?. *Bulletin of Economic Research*, *51*(2), 125-150.

Chen, Y., Dietl, H. M., Orlowski, J., & Zheng, F. (2019). The effect of investment into European football on the market value of Chinese corporations. *International Journal of Sport Finance*, *14*(4), 249-261.

Chung, J., & Hwang, J. H. (2010). An empirical examination of the parimutuel sports lottery market versus the bookmaker market. *Southern Economic Journal*, 76(4), 884-905.

Courneya K. S., & Carron A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport & Exercise Psychology*, *14*(1), 13-27.

Crowder, M., Dixon, M., Ledford, A., & Robinson, M. (2002). Dynamic modelling and prediction of English Football League matches for betting. *Journal of the Royal Statistical Society Series D: The Statistician*, *51*(2), 157-168.

Dixon, M. J., & Coles, S. G. (1997). Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society, 46*(2), 265-280.

Dobson, S., Gerrard, B., & Howe, S. (2000). The determination of transfer fees in English nonleague football. *Applied Economics*. *32*(9), 1145-1152.

Durán, G. A., Guajardo, M., López, A. F., Marenco, J., & Zamorano, G. A. (2021). Scheduling multiple sports leagues with travel distance fairness: An application to Argentinean youth football. *INFORMS Journal on Applied Analytics, 51*(2), 136-149.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417.

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance*, *46*(5), 1575-1617.

Frick, B. (2011). Performance, salaries and contract length: empirical evidence from German soccer. *International Journal of Sport Finance*, *6*(2), 87-118.

Galloppo, G., & Boido, C. (2020). How much is a goal in the football championship worth? Match results and stock price reaction. *International Journal of Sport Finance*, *15*(2), 83-92.

Gerhards, J., & Mutz, M. (2017). Who wins the championship? Market value and team composition as predictors of success in the top European football leagues. *European Societies, 19*(3), 223-242.

Gerrard, B., & Dobson, S. (2000). The determination of transfer fees in English nonleague football. *Applied Economics*, *32*(9), 1145-1152.

Gilbert, E., Dixon, J., & Loughead, T. (2020). An examination of travel effects on performance outcomes in major league soccer. *Journal of Applied Sport Management, 12*(1), 46-59.

Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of forecasting*, *21*(2), 331-340.

Goddard, J., & Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, *23*(1), 51-66.

Goller, D., Knaus, M. C., Lechner, M., & Okasa, G. (2021). 22. Predicting match outcomes in football by an Ordered Forest estimator. *A modern guide to sports economics*, 335-355.

Gövdeli, T., & Güngör, A. Y. (2022). Predictors of crowd effect in football: Evidence from five major football leagues of Europe. *Turkish Journal of Sport and Exercise, 24*(1), 30-37.

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Güngör, A. Y. (2024). Ranking the Performances of Goalkeepers by WASPAS Method: AFC Asian Cup Qatar 2023. *Spor ve Rekreasyon Araştırmaları Dergisi, 6*(1), 25-35.

Hagen, J., & Cunha, M. N. (2019). The History of Investing in Football and Factors Affecting Stock Price of Listed Football Clubs. *International Journal of Financial Management*, *9*(4), 31.

Holmes, B., & McHale, I. G. (2023). Forecasting football match results using a player rating based model. *International Journal of Forecasting*, *40*(1), 302-312.

Hubáček, O., Šourek, G., & Železný, F. (2019). Exploiting sports-betting market using machine learning. *International Journal of Forecasting, 35*(2), 783-796.

Internet – Grand View Research – Sports Betting Market Size, Share & Trends Analysis Report By Platform, By Betting Type (Fixed Odds Wagering, Exchange Betting, Live/In-Play Betting, eSports Betting), By Sports Type, By Region, And Segment Forecasts, 2023 – 2030. (2023). Retrieved from https://www.grandviewresearch.com/industry-analysis/sports-betting-market-report#:~:text=The%20sports%20betting%20market%20was,USD%20182.12 %20billion%20by%202030

Internet – no author, no date – Retrieved from https://www.oddsportal.com/football/poland/ekstraklasa/

Internet – no author, no date – Retrieved from https://www.oddsportal.com/football/turkey/super-lig/

Internet – no author, no date – Retrieved from https://www.transfermarkt.com/super-lig/gesamtspielplan/wettbewerb/TR1/saison_id/2021

Internet – no author, no date – Retrieved from https://www.transfermarkt.com/pko-bp-ekstraklasa/gesamtspielplan/wettbewerb/PL1/saison_id/2021

Internet – no author, no date – Retrieved from www.erasmus-plus.ec.europa.eu/resources-and-tools/distance-calculator

Jedraszka, A., & Zaton, M. (2011). Giełdy zakładów sportowych jako alternatywne źródło inwestycji, *Studia i Prace Kolegium Zarządzania i Finansów*, 106, 159-170

Kesenne S. (2015). Revenue sharing and absolute league quality; Talent investment and talent allocation. *Scottish Journal of Political Economy*, *62*(1), 51-58.

Koopman, S. J., & Lit, R. (2019). Forecasting football match results in national league competitions using score-driven time series models. *International Journal of Forecasting*, *35*(2), 797-809.

Lucifora, C., & Simmons, R. (2003). Superstar effects in sport. Evidence from Italian soccer. *Journal of Sports Economics*, *4*(1), 35-55.

Majewski S. (2016). Identification of factors determining market value of the most valuable football players, *Journal of Management and Business Administration. Central Europe*, *24*(3), 91-104.

Majewski S., & Majewska A. (2018). Econometric modelling of market prices of football clubs 'sponsors' stocks, *The IEB International Journal of Finance*, 17, 2-17.

Majewski, S. (2014a). Modelling of football companies' rates of return according to sport results and bookmakers' expectations on the example of Serie A. *Business and Economic Horizons*, *10*(3), 214-222.

Majewski, S. (2014b). Szacowanie wartości rynkowej piłkarskich kart zawodniczych przy wykorzystaniu modeli ekonometrycznych. *Finanse, Rynki Finansowe, Ubezpieczenia*, 66, 663–674.

Malagón-Selma, P., Debón, A., & Domenech, J. (2023). Measuring the popularity of football players with Google Trends. *Plos One*, *18*(8), 1-21

McCarthy, J., Rowley, J., Jane Ashworth, C., & Pioch, E. (2014). Managing brand presence through social media: the case of UK football clubs. *Internet Research*, *24*(2), 181-204.

Metelski, A., Kornakov, K., Sharda, R., & Chłapowski, J. (2022). How the value of football players influences a team's chances of victory-a Euro 2020 example. *Journal of Physical Education and Sport, 22*(1), 167-173.

Sauer, R. D. (1998). The economics of wagering markets. *Journal of Economic Literature*, *36*(4), 2021-2064.

Škrinjarić, T., & Barišić, P. (2019). Effects of football match results of Croatian national team on stock returns: Evidence from Zagreb stock exchange. *Zagreb International Review of Economics & Business*, *22*(1), 13-45.

Snyder, W. W. (1978). Horse racing: Testing the efficient markets model. *The Journal of Finance*, *33*(4), 1109-1118.

Sóti, A., Ancarani, A., Stéger, J., & Vattay, G. (2020). Influence of Twitter activity on the stock price of soccer clubs. *Social Network Analysis and Mining*, 10, 1-12.

Szymanski, S., & Weimar, D. (2019). Insolvencies in professional football: A German Sonderweg. *International Journal of Sport Finance*, *14*(1), 54-68.

Thaler, R. H., & Ziemba, W. T. (1988). Anomalies: Parimutuel betting markets: Racetracks and lotteries. *Journal of Economic Perspectives*, *2*(2), 161-174.

Trequattrini, R., Lombardi, R., & Nappo, F. (2012). The evaluation of the economic value of long lasting professional football player performance rights. *WSEAS Transactions on Business and Economics*, *9*(4), 199-218.

Tunaru, R., Clark, E., & Viney, H. (2005). An option pricing framework for valuation of football players. *Review of Financial Economics*, *14*(3-4), 281–295.

Warias, R., & Block, S. (2021). Influence of the Team Value of Professional Football Clubs on Their Success in Sports. In *SHS Web of Conferences* (Vol. 129, p. 03035). EDP Sciences.

Wicker P, Prinz J., Weimar D., & Upmann T. (2013). No pain no gain? Effort and productivity in professional soccer, *International Journal of Sport Finance*, *8*(2), 124-139.

Williams, L. V. (1999). Information efficiency in betting markets: A survey. *Bulletin of Economic Research*,