Stockholm School of Economics
Department of Accounting
Bachelor Thesis in Accounting & Financial Management
May 2022

Valuation in English Football

The Relationship Between Sports Performance and Financial Valuation of Football Clubs

Andreas Rylander
24821@student.hhs.se

Lucas Ternqvist
24730@student.hhs.se

Abstract

This paper aims to contribute to the field of accounting and sports by investigating the relationship between sports performance and financial valuation in English football, using a sample of 72 football clubs in the English Premier League and the EFL Championship from 1998 to 2019. Markham's multivariate method is applied, a valuation model based on accounting metrics, to determine football clubs' approximate and comparable financial valuations for a specific year. Using regression analysis, obtained results suggest a relationship between sports performance and financial valuation exists. In particular, an exponential relationship is established for the Premier League. In addition, the validity of Markham's multivariate method is examined by comparing it to the market capitalization of Manchester United, the only publicly listed club in English football.

Supervisor: Katerina Hellström

Keywords: Financial valuation, Sports performance, Markham's multivariate method, Football industry, Exponential regression

Acknowledgements: We want to express our deepest gratitude to our supervisor Katerina Hellström for her guidance during the course of this project. Also, we would like to give a special thanks to Oliver Lindqvist Parbratt for his support and for providing us with the data set that made this work possible.

Contents

1 Introduction

1.1 Background

Football is considered the biggest and most popular sport globally, with an estimated fan base of 4 billion people (Wood 2008; Shvili 2020). Affecting the daily life of billions, football's influence and reach into countries and cultures is unequaled among sports (Nielsen 2018). The modern-day variant of the game has its origins in mid-19th century England, and in 2020 the European football market generated EUR 25.2 billion in revenues. In 2021, Manchester United was considered the most valuable English football club with an estimated brand value of USD 1,327 million, closely followed by Manchester City at USD 1,313 million (Lange 2022).

In particular, English football is considered by many the highest performing and most valuable domestic football industry for several reasons. Firstly, the English Football Association and its domestic leagues have been superior in terms of sports performance in relation to their European counterparts during the past two seasons according to the Country Coefficients determined by UEFA, the governing body of European football (UEFA 2022). Secondly, the country's first division, the English Premier League, is the most viewed football league with an estimated global audience of 3.2 billion for the 2019/2020 season, roughly double that of the second-placed European competition UEFA Champions League and the third-placed German top division Bundesliga (EY 2022). Thirdly, the English Premier League is the world-leading football league concerning financial resources, with 10 clubs making it to the top 20 revenue generating teams worldwide (Deloitte 2022), contributing GBP 7.6 billion to the UK economy in terms of GVA (Gross Value Added, the incremental value added to the country's GDP) (EY 2022). As a result, the football industry poses an interesting area of study within the field of accounting and sports, especially the English football industry. Arguably, researching the largest domestic football industry could result in the greatest impact and most insightful contributions made to the literature.

Based on an extensive literature review (see Section 2.1), targeting the relationship between the financial valuation of football clubs and sports performance is a subject requiring further investigation. From a large body of research studying financial performance in relation to sports performance for football clubs, investigating financial valuation instead of the former would contribute to the field of accounting and sports. If a relationship exists where sports performance can explain the financial valuation of football clubs, this creates multiple incentives for shareholders to improve sports performance. The benefits for club owners would be two-fold: achieving sports success and increased financial wealth from an enhanced club valuation.

1.2 Objectives

The purpose of this thesis is to investigate the relationship between the financial valuation of football clubs and their sports performance. Additionally, it aims to, in practice, utilize the multivariate method for approximate financial valuations of football clubs developed by Markham (2013). The subject is of interest to the academic world of sports accounting and practitioners, especially shareholders, in the football industry. While a body of research proposes contradicting views on how sports performance can explain financial performance, there is limited research on its relation to financial valuation. Hence, valuable contributions are made to the field of sports accounting regarding whether or not sports performance can explain financial valuation. Insights concerning this relationship can also provide information to practitioners and shareholders in terms of extended motives to achieve increased sports performance. Therefore, this thesis aims to answer the following **research question**:

Can the financial valuation of football clubs be explained by their sports performance?

1.3 Limitations

In this study, a data set for English football clubs in the country's first and second division, the English Premier League and the EFL Championship, respectively, will be considered. The data set consists of financial data for 72 clubs in total, including financial key ratios, income statements, and balance sheets for each club from the years 1998 to 2019. Due to the use of a promotion and relegation system, the observed clubs differ over time, and the data set covers only 44 clubs yearly, 20 clubs in the Premier League and 24 clubs in the EFL Championship. Since only Manchester United is publicly listed among English football clubs, this study almost exclusively concerns football clubs' valuation in a non-public setting. Consequently, club values are approximations and not "true" market values. Hence, from a valuation perspective, this study is limited to the valuation method developed by Markham (2013), which applies to any football club. Furthermore, due to few football clubs being publicly listed globally, the reliability of deriving market values as industry benchmarks is deemed unsatisfactory and is therefore not considered.

1.4 Outline

This study is divided into five additional sections. In Section 2, a literature review including relevant theoretical background is provided. Following this, in Section 3, the data set and the methodology is presented. In Section 4, all results are presented and analyzed concerning the research question. Section 5 discusses the obtained results with respect to different considerations and limitations. Lastly, in Section 6, conclusions are given based on the results and the discussion.

2 Literature Review and Theory

This section is divided into four parts. Firstly, a literature review concerning the research topic is provided. Secondly, relevant theoretical frameworks related to accounting and financial valuation are presented, in particular conventional company valuation methods. Thirdly, the valuation of football clubs is covered regarding the implications of valuation in the football industry, and an alternative and approximate approach for club valuation is presented. Lastly, the developed hypothesis to be tested in this thesis is formulated.

2.1 Literature Review

Empirical capital market research in accounting evolved in the 1960s, with the paper by Ball et al. (1968) being recognized by many as one of the first articles in the field (Karuna 2019). In their paper, Ball et al. (1968) manage to establish a link between accounting information, in terms of net earnings, and stock prices. The development of the field eventually led to a wave of research centered around fundamental analysis and accounting-based valuation in the 1980s and 1990s, for instance, through the residual income model proposed by Ohlson (1995) and Feltham et al. (1995) (Skogsvik 2002). Simultaneously, the presence and value of sports have increased expeditiously during recent decades, leading to an important cultural and economic phenomenon in many parts of the world. The increased commercial value of sports has given rise to a growing body of literature on the intersection of accounting and sports (Andon et al. 2019).

Like professional sports overall, the football industry has seen rapid growth in recent decades, with the English football industry positioned as the domestic football industry with the largest international appeal. This is reflected in the Premier League’s broadcasting exports comprising almost half of the country’s television exports, consistently exceeding the combined total of the Spanish (La Liga), German (Bundesliga), Italian (Serie A), and French (Ligue 1) domestic top divisions in the last decade. However, as sports in general and English football in particular is positioned in the union of commercial fields and popular culture, there are some important characteristics to consider that differentiate it from other industries and businesses (Cooper et al. 2012), such as the objectives of a football club.

Identifying the correct objectives of football clubs has always been subject to controversy, whether it regards only sports performance or if it also includes financial performance. The football industry is a well-defined market consisting of different clubs, much like competing specialist firms, but to achieve sports success. Nevertheless, football clubs seldom make money except for the top clubs. Hence, the football industry is an example of a mature market, although some specific characteristics suggest that profit maximization is not the primary objective. This is exemplified by the fact that loss-making clubs are not likely to exit the market, as firms generating non-negative profits in other industries would (Szymanski et al. 1997).

For instance, Dimitropoulos (2014) suggests that football clubs operate near the edge of financial distress but with a low possibility of bankruptcy since multiple stakeholders are committed to bailing the clubs out if needed. Another explanation is the limited corporate control in the football industry. This stems from a lack of pressure from financial markets since most football clubs are not publicly listed, and that football club owners' main concerns relate to sports performance (Szymanski et al. 1997). Consistent with this, Sloane (1971) studied football clubs from a utility-maximizing perspective and argues that their aim is, while remaining solvent, to achieve sports success. As a result, generating profits is a secondary objective supporting the main priority of playing football and winning games. Malagila et al. (2021) also support this while further stating that football clubs tend to prioritize on-the-field performance (sports performance) at the cost of off-the-field performance (financial performance). This further strengthens the thesis that football clubs focus on sports performance over financial performance, at least to the point of financial distress. From a sports perspective, this makes sense as football clubs originally, in almost every case, are established as member associations with non-profit objectives and with the purpose of playing football (Peeters et al. 2013).

Since football is one of the biggest and most popular sports worldwide, football clubs have been widely studied in relation to the field of accounting. In earlier research, numerous studies have investigated the relationship between financial performance and sports performance of football clubs. However, there are contradicting views on this matter. Simone et al. (2020) suggest that a significant and one-way relationship exists, where financial performance positively impacts sports performance, while Alaminos et al. (2020) state that financial performance is dependent on sports performance. The authors further state that liquidity and leverage also are significant factors explaining financial performance. Additionally, on some occasions, it is stated that the reputation of a football club also impacts its financial performance. On the contrary, other research suggests no relationship between sports performance and financial performance exists, and Galariotis et al. (2018) argue that higher revenue and better sports performance do not aid financial performance due to short-term and sports-related objectives. As a result, the long-term financial performance of football clubs suffers. Szymanski (1998) also confirms this by stating that no systematic relationship between profits and league performance exists.

Moreover, multiple studies suggest that sports performance of football clubs is dependent on wages, implying that paying higher salaries will result in better performance of the team. This indicates that a higher wage level enables owning better players that can achieve better sports results (Ferri et al. 2017; Pereira 2018; Szymanski 1998). Furthermore, the literature suggests that there is a direct relationship between revenues and sports performance for football clubs, as increased revenues results in better sports performance (Pereira 2018; Galariotis et al. 2018; Szymanski 1998). This implies that generating more revenue enables further investments in the team to improve sports performance. Contrariwise, better sports performance will generate higher revenue since both fans and sponsors are attracted to a team's success (Szymanski 1998).

In contrast to investigating the relationship between financial performance and sports performance of football clubs, there is limited research on the relationship between the financial valuation of football clubs and their sports performance. This may be due to certain implications of financial valuation attributed to specific characteristics prevailing in the industry, including the non-profit objective of football clubs. However, Klobučník et al. (2019) aim to investigate the relationship between football clubs' sports performance, their market value, and the economic performance of their geographical region.

The authors conclude that a strong positive correlation between the market value of football clubs and their sporting success exists at a regional level. Furthermore, the study identified a moderate correlation between a geographic region's economic performance (measured in GDP) and the market value of football clubs. It is stated that, in general, clubs with higher market values are located in countries with high economic performance, indicating that population, territory, quality of the national football league and game attendance drive the market value of football clubs.

In a listed setting for a few Italian football clubs, Botoc et al. (2019) find that a club's share price is sensitive to positive football results. Nevertheless, this is based on a short-term perspective from investigating daily stock prices in relation to single football games. While these results imply that sports performance positively affects a club's valuation, the same conclusions may not be drawn for clubs in an unlisted environment lacking a market value. In addition, there are limitations to drawing conclusions regarding valuation from a long-term perspective since targeting single games and daily stock price changes do not guarantee similar outcomes in the long run.

Furthermore, valuation in the football industry comes with implications that are non-existent in other sectors. Tiscini et al. (2016) argue that football clubs should not be valued from a stand-alone perspective, which is the common approach to value companies (Koller et al. 2015). This differs from other sectors, and empirical evidence shows that, on average, football clubs yield negative economic results, yet the difference in market value and book value is generally positive. Hence, the market value of a football club is likely driven by other factors. Consequently, conventional valuation methods, such as discounted cash flow methods and trading multiples, are not appropriate. Furthermore, Tiscini et al. (2016) state that the value of a football club is driven by sales turnover, rather than net income, as well as overall shareholder benefits, including private benefits of control and socio-emotional benefits tied to the reputation of the club and its sports performance. Due to this, profitability in the football sector does not drive valuation to the same extent as it would in other industries, and the club's brand image is a more impactful value driver. Brand image can also drive revenue and strengthen a club's market position. For instance, Manchester United has historically dominated the market with higher ticket prices despite poorer league performance than rivals while also attracting higher attendance (Szymanski 1998).

Moreover, Tiscini et al. (2016) state that, due to a general lack of profitability in the football industry, equity capital opportunities are substantially zero. Still, investors are willing to invest in football clubs due to external or emotional private benefits since controlling shareholders enjoy other remunerations from their investments. The authors further state that in order to capture the value of a football club, a revalued net assets value method is deemed more appropriate than conventional valuation methods, especially by taking into account the fair value of intangible assets. In the football industry, player registration rights are recognized as intangible assets (PWC 2018) and represent the players' monetary value. Furthermore, the main valuation drivers are identified as sports success in combination with reputation, where sports performance is naturally linked to specific players representing the club and reputation, and brand value is linked to commercial revenues (Tiscini et al. 2016).

From a traditional investor's approach, including long institutional investors, there is limited attractiveness to investments in football clubs. Historically, football stock indices show erratic behavior, likely due to high uncertainty about a club's fair value. Since the business model and balance sheet are closely linked to sports performance, this has implications for valuation purposes (Aglietta et al. 2010). According to Dalso et al. (2020), there are implications in finding a general valuation approach targeting football clubs. Due to different biases, the valuation technique must be tailored to deliver satisfying results. However, a revenue multiple valuation shows promising results in capturing actual demand in contrast to using a strict risk-return approach, as with discounted cash flow methods. Therefore, multiple valuation is a better method in general, even though it is not optimal. Nonetheless, there are also limitations to trading multiples, as most clubs are not publicly listed. Consequently, establishing an adequate peer group is not easy since market values can only be derived from a few listed clubs.

New methods for valuation of football clubs, or at least for approximating club values, have been evaluated. For instance, Markham (2013) introduced a multivariate method to value football clubs with the aim of creating a universal approach applicable to any football club, listed or not. In his paper, the model was tested on a sample of Premier League clubs between the years 2004-2012 alongside several different valuation methods, where his method provided the most consistent and reliable results. The model was developed as an alternative method to other established valuation techniques, such as discounted cash flow methods and trading multiples that are not reliable and universally applicable in this context. The approach is based on a multiplier model, being an extension of using a revenue multiple that is solely considered too simplistic, with the aim to capture a sports franchisee's ability to generate future revenue. As this method remains relatively unexplored, the author suggests that further research may be conducted to fit a similar model to different leagues, for instance, the second division in English football, namely the English Football League Championship. However, due to the lack of listed football clubs providing information on market values, Markham's valuation method may be seen as an approach for approximation. Moreover, it can provide comparable valuation estimates for all football clubs based on identified value drivers in the football industry.

Following the introduction of Markham's multivariate method, suggestions have been made to extend the model through minor modifications. In order to capture large transfer values, net income is proposed to be substituted with EBITDA, while amortization costs can be added to the wage ratio. This slightly modified version produces similar results to the original model, but it rewards clubs that spend less in the transfer market than other clubs with similar size, revenue, and profit (The Esk 2017).

2.2 Conventional Company Valuation Methods

According to classical theory of asset pricing, the fair price of a financial asset equals the sum of expected future cash flows discounted by the cost of capital (Fabozzi et al. 2017). Hence, what drives the value of a company is its ability to generate future cash to investors. From the law of one price, by discounting future cash flows, the present value represents the amount of cash today required to generate the same future cash flows from investing elsewhere with the same risk (Berk et al. 2016). Moreover, surveys show that the two main approaches used in practice for equity valuation are a market-based approach through trading multiples and a fundamental approach through discounted cash flow valuation (Pinto et al. 2019; Demirakos et al. 2004). While discounted cash flow valuation is a more ”sophisticated” approach, it comes with certain limitations regarding technical applicability, which causes analysts to rely on market-based valuation multiples. By taking market values of comparable firms into account, multiples provide a valuation that ”feels right” (Imam et al. 2008). Consequently, both approaches are often combined, as discounted cash flow valuation is deemed more accurate, while trading multiples can provide insights for summarizing and testing valuation results (Koller et al. 2015; Imam et al. 2008).

Enterprise discounted cash flow (DCF) valuation discounts free cash flow (FCF), which represents the value generated by the firm to all types of investors (equity and debt holders, as well as other types of investors) at the weighted average cost of capital (WACC), the blended cost of capital for all types of investor capital (Koller et al. 2015). This is used to determine the firm's equity value by subtracting the firm's debt and other non-equity claims on the FCF from the enterprise value. Often, the FCFs are determined using a five to ten year forecasting period and a terminal value to capture all of the future value generated by the firm. This enables incorporating specific information about a firm's financial outlooks and tends to be more accurate than trading multiples (Berk et al. 2016). Moreover, FCF is mainly driven by the expected return on invested capital (ROIC) and revenue growth. In essence, a firm creates value as long as the ROIC exceeds the WACC. Given that DCF valuation relies on a firm's in- and outflows of cash instead of accounting-based earnings, it remains a popular valuation method among academics and practitioners (Koller et al. 2015).

Trading multiples may be seen as a "short-cut" version of DCF valuation. Instead of conducting fundamental analysis of the firm, a valuation is derived from the market's assessment of similar prospects, often referred to as peers. The method provides simplicity and is based on actual market values rather than forecasts of future FCFs. Common multiples are the Price-to-Earnings ratio (P/E) or various enterprise value multiples (e.g. EV/EBITDA). However, the main limitation of this approach is to find comparable firms to establish a relevant peer group (Berk et al. 2016). While a DCF approach is useful for providing a detailed multi-period forecast, a trading multiples approach is more suitable from a short-term perspective (Imam et al. 2008).

2.3 Valuation of Football Clubs

DCF valuation is especially useful for multi-business companies and works best when applied to projects, business units, and companies with a consistent capital structure (Koller et al. 2015). Since football clubs cannot be identified as multi-businesses and often have deteriorating capital structures from increased debt levels (Dimitropoulos 2014), a DCF approach is likely not appropriate. Furthermore, applying DCF valuation to football clubs is questionable due to the lacking profit objective since profitability is one of the main valuation drivers in this method. As the literature suggests (see Section 2.1), specific characteristics of the football industry result in other main value drivers such as turnover, brand image, and reputation rather than the factors considered in conventional valuation methods. Moreover, using trading multiples to value football clubs may not be a suitable approach either, mainly due to the lack of publicly listed clubs. Hence, it is hard to establish a peer group to derive market value benchmarks. However, if using this approach, the research indicates (see Section 2.1) that a revenue multiple is more appropriate than multiples based on profitability metrics.

Due to the implications of valuing football clubs, the reliability of conventional valuation methods can be questioned. With the aim of creating a universal approach, Markham (2013) developed an alternative approach to determine approximate and comparable valuations of football clubs, resulting in the following multivariate method

$$Club\ Valuation = (Revenue + Net\ Assets)\left(\frac{Net\ Profit + Revenue}{Revenue}\right)\left(\frac{Stadium\ Capacity\ \%}{Wage\ Ratio\ \%}\right), \tag{1}$$

where the first term consists of revenue, representing the total revenue generated by the club in a financial year, and net assets, i.e. total assets less total liabilities. This term makes up the backbone of the valuation model and demonstrates the club's ability to generate future revenue. It is then multiplied by net profit (or loss) added to revenue, divided by revenue. This figure represents profitability in relation to overall revenue. Consequently, this factor will be less than one for loss-making clubs and greater than one for clubs making a profit. Finally, the overall figure is multiplied by stadium capacity divided by the wage ratio, where stadium capacity is the average stadium utilization illustrating a club's ability to utilize its core asset. The wage ratio is defined as wages divided by revenue, describing a club's ability to control its major expenditures in terms of salaries (Markham 2013).

2.4 Hypothesis Development

Based on the literature review (see Section 2.1), a few insights are made that underpin the developed hypothesis.

Firstly, research shows contradicting findings regarding the relationship between the financial performance of football clubs and their sports performance. However, there is limited research on the relationship between financial valuation and sports performance, making this an interesting topic for further investigation.

Secondly, financial valuation in the football industry differs from other sectors, as DCF analyses and trading multiples are not appropriate due to specific characteristics that influence valuation. Here, factors such as sports performance and brand image drive valuation more than profitability. Moreover, the multivariate method developed by Markham (2013), applying to both listed and unlisted clubs, is an alternative method to value football clubs. This method considers several different value drivers, such as revenue, net assets, profitability, home game attendance, and wages.

Thirdly, revenue is positively correlated with the sports performance of football clubs, and it is found that sports performance is reliant on wages. With revenue being the main valuation driver for football clubs, according to Markham (2013), there are reasons to believe that sports performance and financial valuation are positively correlated. However, as sports performance is dependent on wages, and increased wages decrease club valuation using this method, this poses an interesting aspect of the relationship between sports performance and financial valuation. As the main valuation driver, revenue will possibly offset the negative influence that wages have on valuation.

Based on these findings, it is likely that sports performance can explain the financial valuation of football clubs. Hence, the null hypothesis to be tested in this thesis is formulated as:

H_0: *There is no positive correlation between a football club's sports performance and its valuation*

3 Data and Methodology

This section is divided into four parts. Firstly, the data set used to conduct this study is presented. Secondly, the methodology used to test the hypothesis stated in Section 2.4 is described, including the different regression models used. Thirdly, the variables used in the regression models are introduced. Finally, statistical considerations are covered.

3.1 Data set

The data set was received from the *Center for Sports and Business* at *Stockholm School of Economics* and consists of data extracted from financial reports for English football clubs in the top two divisions, comprised of the English Premier League (Premier League) with 20 teams and the English Football League Championship (The Championship) with 24 teams, for the years 1998-2019. In addition to line items pertaining to the income statement and the balance sheet, the data set includes sports performance metrics such as league position, total points, and goal difference for each team and year.

As seen in Section 2.3, using Markham's multivariate method, an approximated club value can be computed depending on accounting line items found in or derived from financial reports. However, stadium capacity (defined as average stadium utilization) was not included in the original data set and had to be collected manually for each club and year (Transfermarkt 2022; Worldfootball 2022). In Table 3.1, all items required to calculate club value are defined.

Item	Name	Definition
Revenue	RE	Total revenue generated for the year
Net Assets	NA	Total assets less total liabilities
Net Profit	NP	Total profit (or loss) for the year
Stadium Capacity	SC	Average stadium utilization $(\frac{Average\ attendence}{Stadium\ capacity})$
Wage Ratio	WR	Wages in relation to revenue $(\frac{Wages}{Revenue})$
Club Value	CV	Approximated valuation of club

Table 3.1: Definition of items in Markham's multivariate method.

In order to analyze the Premier League and The Championship in combination, as well as both leagues separately, the original data set has been used to create the following three sets of data

$$Universal\ set = ALL = \{i : i = 1, 2, ..., 44\}, \tag{2}$$

$$EPL = \{i : i = 1, 2, ..., 20\}, \tag{3}$$

$$EFL = EPL^{\complement} = \{i : i = 21, 22, ..., 44\}, \tag{4}$$

where i refers to a club's total league position at the end of each season for each year $t = 1, 2, ..., 22$. The universal set, referring to the original data set, is equal to the set ALL and is used to investigate a combination of both leagues, while the smaller sets EPL and EFL are created to investigate the Premier League and The Championship separately. Henceforth, ALL, EPL, and EFL are simply referred to as data sets.

The original data set consists of 968 data points for the Premier League and The Championship. However, data points resulting in a negative club value have been excluded since it is assumed that club value only can take on non-negative numbers. Based on Equation 1, this can occur in two different scenarios. Firstly, if net assets are negative and larger than revenue in absolute terms, the first component in the equation takes on a negative sign. Secondly, if net profit is negative and larger than revenue in absolute terms, the second component in the equation takes on a negative sign. In both cases, this yields a negative club value. However, if these scenarios occur simultaneously, the yielded club value is positive. This can be considered a flaw in Markham's multivariate method, as it generates a misleading club value, and such data points are also excluded.

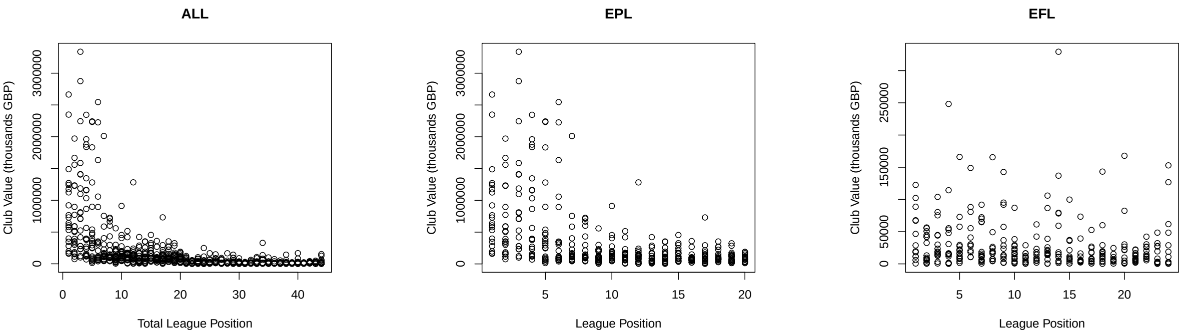
In total, 113 data points resulting in a negative club value have been excluded, of which 20 are from the Premier League, and 93 are from The Championship. For the Premier League, 17 data points resulted in a negative club value from net assets being negative and exceeding revenue in absolute terms, and 3 data points due to a negative net profit exceeding revenue in absolute terms. Likewise, for The Championship, 78 data points resulted in a negative club value from net assets being negative and exceeding revenue in absolute terms, and 15 data points due to a negative net profit exceeding revenue in absolute terms. Additionally, 33 data points are excluded where both scenarios occur concurrently, resulting in a misleading positive club value, of which 3 are for the Premier League and 30 for The Championship. Furthermore, 111 additional data points have been excluded due to missing line items in the original data set, resulting in non-existing club values. This results in 711 observations for both leagues, of which 386 are for the Premier League, and 325 are for The Championship.

In Table 3.2, descriptive statistics are given for the three data sets. The computed club values for teams in the Premier League are significantly higher than those in The Championship. This may be explained by the greater ability of Premier League clubs to generate higher revenue and attract higher attendance, likely due to playing in the top division. Moreover, in line with the literature (see Section 2.1), profits are negative on average for both leagues. This supports the idea that football clubs lack profit objectives and instead prioritize sports performance. Furthermore, note that stadium capacity can take on values larger than one, indicating that game attendance can exceed the formal number of seats in a stadium.

Item	Obs	Min	1st Qu.	Median	Mean	3rd Qu.	Max
ALL							
RE	711	2,972	16,924	40,029	75,972	93,387	608,141
NA	711	−128,465	−4,719	6,681	33,285	32,228	746,642
NP	711	−121,300	−7,287	−1,765	−2,478	2,226	163,930
SC	711	0.091	0.72	0.84	0.82	0.96	1.088
WR	711	0.26	0.51	0.61	0.63	0.73	1.54
CV	711	119	14,850	60,002	198,473	155,409	3,337,656
EPL							
RE	386	14,288	48,954	85,693	122,355	145,705	608,141
NA	386	−128,465	−5,427	21,454	56,952	59,516	746,642
NP	386	−121,300	−10,663	−29	−2,036	6,790	146,419
SC	386	0.091	0.87	0.95	0.92	0.98	1.088
WR	386	0.26	0.49	0.56	0.58	0.66	0.99
CV	386	1,826	74,180	135,967	342,762	346,428	3,337,656
EFL							
RE	325	2,972	9,802	15,700	20,883	26,683	127,947
NA	325	−58,854	−4,494	1,966	5,176	9,212	136,960
NP	325	−45,241	−5,551	−2,119	−3,003	111	163,930
SC	325	0.38	0.60	0.70	0.70	0.78	1.018
WR	325	0.33	0.57	0.67	0.70	0.80	1.54
CV	325	119	5,483	14,748	27,103	29,221	329,313

Table 3.2: Descriptive statistics for all data sets. Monetary amounts are given in GBP (thousands).

In Figure 3.1, the three data sets are visualized through scatterplots representing computed club value and sports performance in terms of league position for all years. The patterns suggest that a correlation between club value and sports performance is more likely to exist in the Premier League than in The Championship. For a combination of both leagues and the Premier League, the scatterplots indicate a slight exponential decay rather than a linear relationship. Based on this, it is deemed reasonable to test for both a linear and an exponential relationship between club value and sports performance.



(a) Both leagues (b) Premier League (c) The Championship

Figure 3.1: Scatterplots for Club Value vs League Position for all data sets.

3.2 Methodology

This study investigates the association between sports performance and the financial valuation of English football clubs. A common analytical tool for association studies is regression analysis which tests the relationship between a dependent variable and a set of related predictor variables. This methodology is useful for expressing the relationship between a dependent variable and a set of related predictor variables (Montgomery et al. 2006).

For testing the hypothesis (see Section 2.4), ordinary least squares (OLS) regression is used where both simple and multiple linear regression is applied. Additionally, given the pattern of an exponential decay observed in the scatterplots for both leagues and the Premier League (see Figures 3.1a and 3.1b), a simple exponential regression model is also used. Below, one model for each of the three regression methods is outlined, each comprising a dependent variable, some independent variable/s, and two variables representing fixed effects.

3.2.1 Simple Linear Regression

The first model is constructed as a simple linear regression (SLR) model applied to the three different data sets. The model is defined as follows

$$\textbf{Model 1: } CV_{it} = \beta_0 + \beta_1 LeaguePosition_{it} + \alpha_{it} + \delta_t + \epsilon_{it}, \tag{5}$$

where $LeaguePosition_{it}$ refers to a club's league position at the end of each season, α_{it} represents fixed effects of club popularity, δ_t represents fixed year effects, and ϵ_{it} is the error term. The intercept β_0 and the coefficient β_1 for the independent variable can be estimated by performing OLS regressions.

3.2.2 Simple Exponential Regression

The second model is constructed as a simple exponential regression (SER) model tested for all three data sets. Based on the scatterplots in Figure 3.1, it is deemed appropriate to apply exponential regression since Figures 3.1a and 3.1b for the data sets ALL and EPL illustrate a trend of exponential decay. The model is constructed as follows

$$\textbf{Model 2: } ln(CV_{it}) = \beta_0 + \beta_1 LeaguePosition_{it} + \alpha_{it} + \delta_t + \epsilon_{it}, \tag{6}$$

which also can be re-written as

$$CV_{it} = e^{\beta_0 + \beta_1 LeaguePosition_{it} + \alpha_{it} + \delta_t + \epsilon_{it}}, \tag{7}$$

where $LeaguePosition_{it}$, α_{it}, δ_t, and ϵ_{it} have the same representations as in Model 1. By performing exponential OLS regressions, using the natural logarithm of the dependent variable in Model 1 ($ln(CV_{it})$) as the dependent variable, the intercept β_0 and the coefficient β_1 for the independent variable can be estimated.

3.2.3 Multiple Linear Regression

The third model is constructed as a multiple linear regression (MLR) model and is only applied to the EPL and EFL data sets. The model is defined as follows

$$\begin{aligned}\textbf{Model 3: } CV_{it} = \beta_0 &+ \beta_1 LeaguePosition_{it} + \beta_2 TotalPoints_{it} \\ &+ \beta_3 GoalDifference_{it} + \alpha_{it} + \delta_t + \epsilon_{it},\end{aligned} \tag{8}$$

which is an extension of Model 1 that includes other independent variables measuring sports performance, namely $TotalPoints_{it}$ and $GoalDifference_{it}$. By performing OLS regressions, the intercept β_0 and the coefficients β_1, β_2, and β_3 for the independent variables can be estimated. Model 3 will not be applied to the data set ALL since total points and goal difference are not comparable between the two leagues. For instance, achieving the same amount of points or the same goal difference in the Premier League and The Championship are not equivalent.

3.3 Variables

This part presents the different variables used in all regression models, including the dependent variable, the independent variables, and the variables representing fixed effects.

3.3.1 Dependent Variable

Based on the valuation method developed by Markham (2013), described in Section 2.3, the dependent variable in the regression models is the computed club value (CV), which can be determined for each specific club and year. It is derived from Equation 1, with the abbreviations stated in Table 3.1, as follows

$$CV_{it} = (RE_{it} + NA_{it}) \left(\frac{NP_{it} + RE_{it}}{RE_{it}} \right) \left(\frac{SC_{it}}{WR_{it}} \right), \tag{9}$$

for club $i = 1, 2, ..., 44$ and year $t = 1, 2, ..., 22$.

3.3.2 Independent Variables

The independent variable used in both Model 1 (see Section 3.2.1) and Model 2 (see Section 3.2.2) is $LeaguePosition_{it}$ (LP), as it is recognized as the primary measure of sports performance for football clubs (Pappalardo et al. 2017) and thus serves this purpose well. Furthermore, this measure of sports performance is commonly used in earlier research. See for instance (Szymanski et al. 1997; Szymanski 1998; Simone et al. 2020; Galariotis et al. 2018).

For the data set ALL, $LeaguePosition_{it}$ for club $i = 1, 2, ..., 44$ and year $t = 1, 2, ..., 22$ represents the total league position for a combination of the Premier League and The Championship (as in Equation 2), where finishing first in the Premier League is recognized as observation $i = 1$ and finishing first in The Championship is recognized as observation $i = 21$ since there are 20 teams in the Premier League. For the data sets EPL and EFL, it is defined for club $i = 1, 2, ..., 20$ and $i = 21, 22, ..., 44$ respectively for year $t = 1, 2, ..., 22$, as in Equations 3 and 4. However, when presenting the results for the data set EFL, the observations are numbered as $\{1, 2, ..., 24\}$, although they refer to observations $i = 21, 22, ..., 44$ in the data set ALL.

For Model 3 (see Section 3.2.3), $LeaguePosition_{it}$ is complemented by $TotalPoints_{it}$ (TP)

and $GoalDifference_{it}$ (GD) as two additional independent variables that measure sports performance, where TP refers to the number of accumulated points throughout a season and GD is defined as the total amount of goals produced less goals conceded throughout a season. Total points as a measure of sports performance is commonly used in the literature, see for instance (Barros et al. 2006; Ferri et al. 2017). Since a team's total number of points, in relation to its competitors, determines the league position for a specific year, it is a reasonable measure of sports success. However, the two measures are not equivalent, as league position does not imply the total amount of points. Also, total points is not constant over time in determining league position since the same amount of points could yield different league positions for different years, thus providing a different measure of sports performance.

Furthermore, goal difference is a disparate sports performance measure compared to the two previously mentioned and is less commonly used in earlier literature. Thus, its inclusion does not only allow for a broader view of sports performance but also contributes to the existing literature. The inclusion of goal difference as a sports performance measure is motivated in two ways. Firstly, it is the determining factor of league position if more than one club achieve the same number of total points. Secondly, to some extent, it can describe a football team's playing style. Scoring more goals yields a higher goal difference compared to a club with the same number of goals conceded. This indicates an offensive playing style that may be deemed attractive to fans and players, improving club popularity and stadium attendance, and enabling acquisitions of better players. Overall, including goal difference as a sports performance measure could result in new findings on the football industry topic.

3.3.3 Fixed Effects

For all models, two variables for implementation of fixed effects are included in addition to the independent variable/s described in Section 3.3.2.

The variable α_{it} controls for the popularity of each club by incorporating the level of its fixed effects. Since the popularity of a club is not always linked to sports performance, it is not of interest to neither the research question nor this study's aim. Still, as stated in Section 2.1, it is a variable that could influence club valuation and is appropriate to include in the models to account for fixed effects. The popularity α_{it} is defined as a rank derived from the number of Google Searches for club $i = 1, 2, ..., 44$ and year $t = 1, 2, ..., 22$, where the data has been collected manually using Google Trends to order the clubs based on their popularity in an aggregated table. For instance, the most popular club for one year is assigned the rank 1, and the least popular club is assigned the rank 44. However, there are some limitations to using Google searches to measure popularity. Firstly, since Google Trends data is only available from 2004, observations for earlier years (1998-2003) are assigned their popularity rank from 2004. Secondly, it is assumed that the number of Google searches for a club correlates with its popularity. In general, Google searches could also, to some extent, relate to scandals and other negative publicity.

Additionally, given that the data sets consider 22 years, fixed year effects δ_t is applied to account for omitted variable bias created by heterogeneity in the data that is unobserved and constant over time. Since the club value is a monetary value affected by inflation and the increasing amount of money invested in the football industry (Szymanski et al. 2015), δ_t accounts for this by controlling for common attributes and characteristics for the clubs across time.

3.4 Statistical Considerations

In this section, statistical considerations are made regarding the handling of outliers, testing for heteroscedasticity, and investigation of multicollinearity.

3.4.1 Outliers

Observations that are considerably different from others are considered outliers. i.e. extreme observations. Potential outliers should be investigated and, if relevant, be corrected for since they can have a large impact on the regression results (Montgomery et al. 2006). In particular, it is beneficial to consider both the residual and the leverage of each observation in the data set. This can be achieved by applying Cook's distance, which uses a squared distance measure between the least-squares estimates based on all observations, as well as the estimates obtained when removing some observation (Cook 1977; Cook 1979). Figures A.5-A.7 in the Appendix display Cook's distance plots for detecting outliers in Models 1-3. Evidently, outliers exist for all models and data sets.

Winsorizing is applied at 5% to handle outliers in all models, a commonly used level among researchers (Ghosh et al. 2012). Instead of being excluded from the data set, outliers are converted to percentile values (Reifman et al. 2010). Here, extreme observations below the 5% percentile are upwards adjusted to the percentile value before executing the regression. Likewise, extreme observations exceeding the 95% percentile are downwards adjusted. Consequently, a winsorized model corrected for outliers may result in better goodness of fit.

3.4.2 Heteroscedasticity

A critical assumption of OLS regression is homoscedasticity, i.e. the absence of heteroscedasticity, implying a constant variance of the error term. If homoscedasticity does not exist, the efficiency loss in using OLS regression may be substantial, and the results from the regression may not be valid (Breusch et al. 1979). Hence, it is essential to check whether this assumption holds or not. If heteroscedasticity is present and all observations are non-negative, an appropriate solution is to apply a logarithmic transformation of the dependent variable (Wang et al. 1994). Note that this is done in Model 2 for the SER approach (see Section 3.2.2) by using the natural logarithm of observations as the dependent variable.

From investigating the residual plots in Figures A.8-A.10 in the Appendix for Models 1-3, Model 1 seems to suffer from heteroscedasticity. This is observed from both an increase in variance (see for instance Figure A.8b) and an increase in the square root of the standardized residuals for the fitted values (see for instance Figure A.8b). However, the residual plots for Model 2 look decent, where the variance of residuals and the square root of the standardized residuals are more stable (see for instance Figure A.9b and A.9e) after having transformed the dependent variable. Similar to Model 1, heteroscedasticity may also be present in Model 3 (see Figure A.10).

Model	**Data set**	**Dep. var.**	**Ind. var.**	**BP-stat.**	**p-val.**
Model 1	ALL	CV	TP	319.65	$< 2.2e-16^{***}$
Model 1	EPL	CV	TP	150.85	$2.034e-10^{***}$
Model 1	EFL	CV	TP	76.657	0.05097
Model 2	ALL	$ln(CV)$	TP	125.39	$1.023e-05^{***}$
Model 2	EPL	$ln(CV)$	TP	50.72	0.7084
Model 2	EFL	$ln(CV)$	TP	71.99	0.1025
Model 3	EPL	CV	TP, GD	146.30	$1.41e-09^{***}$
Model 3	EFL	CV	TP, GD	81.81	0.02636^{*}

H_0 : *Constant variance for residuals*
****p<0.001, **p<0.01, *p<0.05*

Table 3.3: Breusch-Pagan tests (Models 1-3).

One way to test for heteroscedasticity is to perform a Breusch-Pagan test, which checks if the variance of the residuals is dependent on the independent variables (Breusch et al. 1979). Here, the null hypothesis states that the residuals are distributed with equal variance (i.e. homoscedasticity). Under H_0, the BP test statistic follows a chi-square distribution. Table 3.3 displays the results for Breusch-Pagan tests for Models 1-3 and all data sets. For Model 1, the null hypothesis is rejected for the data sets ALL and EPL but not for EFL. As expected, this suggests implications of heteroscedasticity for the data sets ALL and EPL since the homoscedasticity assumption is violated. For Model 2, the null hypothesis is not rejected for the data sets EPL and EFL, implying that homoscedasticity is present for both leagues separately. For Model 3, using MLR, the null hypothesis is rejected for both the Premier League and The Championship, indicating that heteroscedasticity is present.

Based on the results from the Breusch-Pagan tests, the assumption of homoscedasticity holds in Model 2 for the data sets EPL and EFL, and in Model 1 for the EFL data set. Hence, in all other cases, robust standard errors will be applied to account for heteroscedasticity. If not, non-robust standard errors may result in severely biased estimators in the presence of heteroscedasticity (Croux et al. 2004).

3.4.3 Multicollinearity

Multicollinearity is a statistical concept describing the occurrence of correlation among two or more independent variables (Newbold et al. 2013). If near-linear dependencies exist among regressors, this indicates problems with multicollinearity that result in large variances and covariance for estimating the regression coefficients (Montgomery et al. 2006). Arguably, the independent variables in Section 3.3.2 correlate from being different measures of sports performance, which all contribute to a team's success. For instance, it is probable that achieving more points and improving goal difference will result in a better league position. Hence, correlation among regressors and multicollinearity will be explored for Model 3 (see Section 3.2.3) to discuss its validity, using correlation matrices and variance inflation factors (VIF).

The correlation matrix is a simple approach to investigate multicollinearity among pairs of regressors (Montgomery et al. 2006), as it indicates if and to what extent multicollinearity is prevalent in the model. Furthermore, analyzing VIFs is useful in an extended approach to detect multicollinearity when more than two regressors are involved, as in the MLR approach used in Model 3 (see Section 3.2.3). The VIF measures the combined effect of the dependencies among regressors, and values between $5 - 10$ indicate that the regression coefficients may be poorly estimated due to multicollinearity. A common approach to dealing with multicollinearity is to exclude certain regressor variables from the model (Montgomery et al. 2006), which is done in Section 4.1.3.

4 Empirical Analysis and Results

This section is divided into three parts. Firstly, the regression results for all models are presented. Secondly, the multivariate method by Markham (2013) is evaluated. Finally, an application of the best fitted model is covered.

4.1 Regression Results

This section presents the results for all regression models described in Section 3.2 for all data sets.

4.1.1 Simple Linear Regression

In Table 4.1, the SLR results for Model 1 are displayed. The independent variable is significant on a 0.1% significance level for the data sets ALL and EPL. However, there is no significance for the EFL data set. This suggests that sports performance can explain club value, at least for a combination of both leagues and for the Premier League in isolation, meaning that the null hypothesis (see Section 2.4) can be rejected in these cases. Note that the coefficients take on a negative sign, suggesting that the correlation with club value is negative. This makes sense since better sports performance leads to an improved league position and a lower observed value for the independent variable. For instance, winning the league is recognized as observation 1 for the independent variable. Due to this, league position is negatively correlated with club value, which indicates that sports performance is positively correlated with financial valuation.

Concerning goodness of fit, a combination of the Premier League and The Championship yields the highest adjusted R-squared value, suggesting that the independent variable explains 67.4% of the dependent variable. Applied to the Premier League, the model suggests that a club's league position at the end of each season can explain 62.4% of its club value. However, the adjusted R-squared value for The Championship is relatively low at 21.6%, suggesting that the model does not fit the estimated club values to the same extent as in the previous two cases.

The regression plots for Model 1 with corresponding scatterplots can be found in the Appendix (see Figure A.1).

Data set	$\boldsymbol{ALL}$	$\boldsymbol{EPL}$	$\boldsymbol{EFL}$
Dep.var.	CV	CV	CV
Ind. var.	LP	LP	LP
Intercept (β_0)	$524,574^{***}$	$553,777^{***}$	$100,052^{***}$
Std.err.	$87,132$	$115,287$	$26,714$
t-val.	6.0205	4.8035	3.7450
p-val.	$2.919e-09$	$2.376e-06$	0.000221
Coefficient (β_1)	$-4,401^{***}$	$-10,747^{***}$	-422
Std.err.	570	$3,012.1$	240
t-val.	-7.7155	-3.5679	-1.7560
p-val.	$4.595e-14$	0.0004135	0.080317
Obs.	711	386	325
R-sq.	0.7039	0.6795	0.3560
Adj. R-sq.	0.6740	0.6238	0.2156
F-stat.	23.6	12.2	2.54
p-val.	$< 2.2e-16$	$< 2.2e-16$	$2.784e-07$

Note: Winsorizing is applied at 5%, Robust standard errors are applied to ALL and EPL
****p<0.001, **p<0.01, *p<0.05*

Table 4.1: Simple linear regression results (Model 1).

4.1.2 Simple Exponential Regression

The results for Model 2, using an SER approach based on the trend of exponential decay observed in the scatterplots (see Figure 3.1a, and 3.1b), are displayed in Table 4.2. The model demonstrates significance for the independent variable on a 5% level for all data sets. In particular, for a combination of both leagues and the Premier League, the independent variable is significant at the 0.1% level. Hence, the null hypothesis (see Section 2.4) can be rejected for all data sets using an SER approach.

The goodness of fit improves with an SER approach (Model 2) compared to an SLR approach (Model 1) for all data sets, except for The Championship. For a combination of both leagues, the adjusted R-squared value exhibits a slight increase from 67.4% to 68.3%. Additionally, a considerable increase in the adjusted R-squared value from 62.4% to 70.5% is seen for the Premier League. This suggests that an exponential model better captures the relationship between sports performance (league position) and financial valuation (club value), especially in the Premier League with an increase of 8.1 percentage points for the adjusted R-squared value. However, an SER approach results in an adjusted R-squared value of 19.4% for The Championship, indicating the worst fit of all models.

In Figure A.2 in the Appendix, regression plots for model 2 with corresponding scatterplots can be found on the form of both Equation 6 and 7.

Data set	ALL	EPL	EFL
Dep.var.	$ln(CV)$	$ln(CV)$	$ln(CV)$
Ind. var.	LP	LP	LP
Intercept (β_0)	12.8827***	12.5774***	12.0159***
Std.err.	0.2310	0.2086	1.3142
t-val.	55.7686	60.3080	9.1430
p-val.	$< 2.2e-16$	$< 2e-16$	$< 2e-16$
Coefficient (β_1)	-0.0685***	-0.0313***	-0.0274*
Std.err.	0.0057	0.0091	0.0118
t-val.	-12.0656	-3.4630	-2.3220
p-val.	$< 2.2e-16$	0.000605	0.0210
Obs.	711	386	325
R-sq.	0.7115	0.7488	0.3380
Adj. R-sq.	0.6825	0.7051	0.1936
F-stat.	24.48	17.15	2.341
p-val.	$< 2.2e-16$	$< 2.2e-16$	$2.69e-06$

Note: Winsorizing is applied at 5%, Robust standard errors are applied to ALL
****p<0.001, **p<0.01, *p<0.05*

Table 4.2: Simple exponential regression results (Model 2).

4.1.3 Multiple Linear Regression

In Model 3, an MLR approach is applied to investigate the relationship between club value and sports performance in the Premier League and The Championship. However, as mentioned in Section 3.4.3, implications of multicollinearity have to be explored due to potential correlation among regressors.

The correlation matrices for the Premier League and The Championship are displayed in Table 4.3, showing pair-wise correlations among all variables. It is established that $TotalPoints$ and $GoalDifference$ are strongly correlated. In addition, both $TotalPoints$ and $GoalDifference$ correlate with $LeaguePosition$. These pair-wise correlations are significant on a 0.1% level and have Pearson correlation coefficients larger than 0.90 in absolute terms.

Table 4.4 presents the VIF estimates for all independent variables, showing alarmingly high values for both the Premier League and The Championship, in particular for $TotalPoints$. Therefore, this variable is excluded from the model, improving the VIF estimates for the remaining variables $LeaguePosition$ and $GoalDifference$ for both data sets. However, the VIF estimates are still in the range of $5 - 10$, suggesting that the problem of multicollinearity may still not be resolved in this reduced model. Consequently, the validity of the MLR results can be questioned. Henceforth, this reduced version of Model 3, excluding $TotalPoints$ as a regressor, will be evaluated in the MLR approach.

$\boldsymbol{EPL}$	(1)	(2)	(3)	(4)
(1) CV	1.00			
(2) $LeaguePosition$	-0.53^{***}	1.00		
(3) $TotalPoints$	0.59^{***}	-0.95^{***}	1.00	
(4) $GoalDifference$	0.58^{***}	-0.91^{***}	0.96^{***}	1.00

$\boldsymbol{EFL}$	(1)	(2)	(3)	(4)
(1) CV	1.00			
(2) $LeaguePosition$	-0.20^{***}	1.00		
(3) $TotalPoints$	0.17^{**}	-0.96^{***}	1.00	
(4) $GoalDifference$	0.19^{***}	-0.90^{***}	0.92^{***}	1.00

****p<0.001, **p<0.01, *p<0.05*

Table 4.3: Pearson correlation matrices (EPL and EFL).

Variable	$\boldsymbol{EPL}$	$\boldsymbol{EPL}$ *(excl. TP)*	$\boldsymbol{EFL}$	$\boldsymbol{EFL}$ *(excl. TP)*
$LeaguePosition$	11.95	7.33	16.01	6.54
$TotalPoints$	27.13	-	20.59	-
$GoalDifference$	17.10	8.41	8.34	6.64

Table 4.4: VIF estimates for the regressors in Model 3 (EPL and EFL).

The MLR results for Model 3, excluding the independent variable $TotalPoints$, are presented in the Appendix (see Table A.1), where the only coefficient showing significance is $GoalDifference$ at a 5% level. However, the implications of multicollinearity may undermine the statistical significance of any independent variable (Montgomery et al. 2006). Hence, there is uncertainty about whether or not the null hypothesis (see Section 2.4) can be rejected for goal difference as a sports performance measure explaining club value due to its high VIF estimate of 6.64. Moreover, since league position shows significance as a sports performance measure explaining club value in both Model 1 and 2 but not for Model 3, this raises additional concerns regarding the reliability of the MLR results. Due to the presence of multicollinearity, also when excluding $TotalPoints$, Model 3 is not deemed an acceptable model. However, the results still provide insights about using $LeaguePosition$ in combination with $GoalDifference$ and $TotalPoints$ as sports performance measures explaining club value, implying that this is not feasible due to multicollinearity.

In Figure A.4 in the Appendix, the regression planes for Model 3 are displayed for both data sets with corresponding three-dimensional scatterplots.

4.1.4 Summary of Regression Results

Three regression models have been applied to the different data sets with varying results from a goodness of fit perspective. A summary of the adjusted R-squared values for all regressions is presented in Table 4.5, where Model 2 (SER) demonstrates the highest adjusted R-squared value for data sets ALL and EPL. Model 1 (SLR) shows slightly better goodness of fit for the EFL data set than Model 2 (SER). An SER approach applied to the Premier League yields the highest adjusted R-squared value of 70.5%, suggesting that this is the best fitted model.

Data set	$\boldsymbol{ALL}$	$\boldsymbol{EPL}$	$\boldsymbol{EFL}$
Model 1	0.6740	0.6238	0.2156
Model 2	0.6825	**0.7051**	0.1936
Model 3	-	0.6286	0.2152

Table 4.5: Adjusted R-squared values (Models 1-3).

4.1.5 Model Evaluation

To investigate the impact of outliers on the goodness of fit, winsorizing is applied for Models 1-3 for all data sets and Table 4.6 displays results for the adjusted R-squared values. As mentioned in Section 3.4.1, the regression results presented in Section 4.1 are based on winsorized models at the 5% level. Up to the 5% winsorizing level, the adjusted R-squared value increases for all models and data sets. For Model 1, the data set ALL yields the highest total absolute increase in the adjusted R-squared value from 53.0% to 73.2% at the winsorizing level 10%, suggesting that outliers may significantly impact the regression results for the SLR approach.

For all data sets, Model 2 demonstrates the smallest total absolute increase in the adjusted R-squared value when adjusting for outliers. However, for the Premier League and The Championship, the adjusted R-squared values slightly decrease from a winsorizing level of 5.0% to 10.0%. For Model 2, this suggests that the 5.0% winsorizing level used in Section 4.1 is appropriate for the Premier League and The Championship. Furthermore, for the Premier League, demonstrating the highest adjusted R-squared value in the original case (0.0% winsorizing level) and at the 5.0% winsorizing level, this strengthens the thesis that an SER approach yields the best fit.

		Winsorizing level				
Model	**Data set**	**0.0%**	**1.0%**	**2.5%**	**5.0%**	**10.0%**
Model 1	ALL	0.5299	0.5604	0.6123	0.6740	0.7321
Model 1	EPL	0.5333	0.5518	0.5672	0.6238	0.6823
Model 1	EFL	0.1483	0.1804	0.1911	0.2156	0.2295
Model 2	ALL	0.6557	0.6610	0.6661	0.6825	0.6822
Model 2	EPL	0.6618	0.6710	0.6914	0.7051	0.7010
Model 2	EFL	0.1860	0.1875	0.1927	0.1936	0.2129
Model 3	EPL	0.5440	0.5608	0.5744	0.6286	0.6821
Model 3	EFL	0.1505	0.1815	0.1920	0.2152	0.2279

Table 4.6: Adjusted R-squared values for Models 1-3 (post winsorizing).

4.2 Evaluation of Markham's Multivariate Method

Throughout this thesis, the multivariate method developed by Markham (2013) is used to compute the financial valuations of football clubs due to them being privately owned entities. However, one football club in the data set is publicly listed, namely Manchester United FC (NYSE:MANU). This constitutes an opportunity to evaluate the validity and reliability of this valuation method by comparing the estimated club value, computed using Markham's multivariate method, and the market capitalization of Manchester United from 2012 to 2019. The market capitalization is derived from the share price at the end of Manchester United's fiscal year, ending on June 30, as the accounting metrics which constitute Markham's club value are given at this point in time.

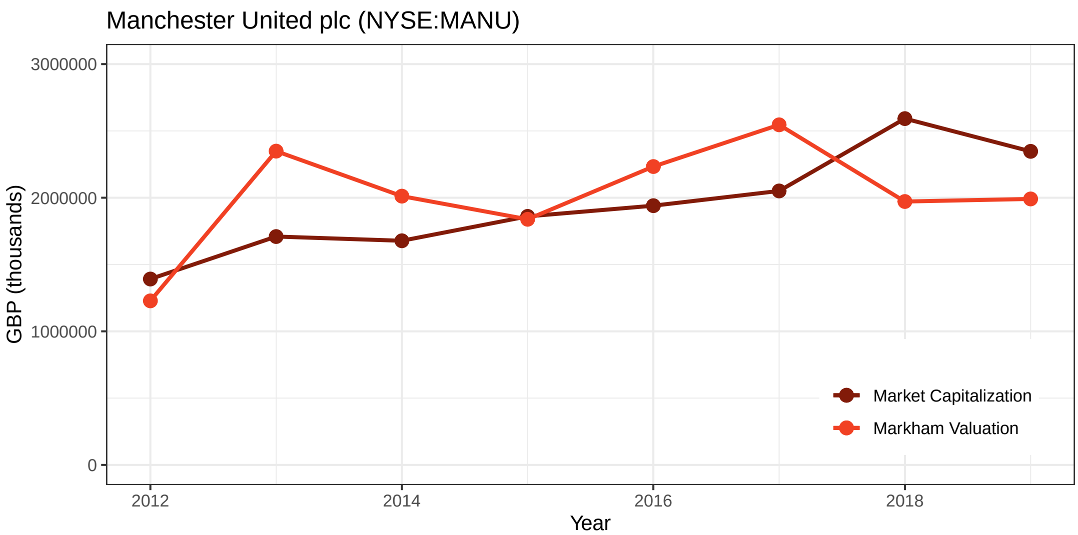


Figure 4.1: Computed club value using Markham's multivariate method (red line) compared to the market capitalization of Manchester United FC (dark red line) for the years 2012-2019.

As seen in Figure 4.1, the club value derived from Markham's multivariate method (labeled Markham valuation) follows the market capitalization of Manchester United fairly well. Worth noting is that Manchester United is among the top five valued clubs using Markham's valuation (see Table A.2 in the Appendix) but is considered the highest valued club in English football (Lange 2022). Although Manchester United's computed club value is not entirely accurate, it still scores among the top clubs and exhibits reasonable estimates compared to its market capitalization. Also, for clubs generating higher Markham valuations, there is no possibility for comparison against ”true” market values.

Year	Share Price (£)	Market Cap.	Markham Valuation	Difference	RT Ratio
2012	8.96	1,391,413	1,227,979	163,434	11.7%
2013	10.43	1,709,299	2,348,595	−639,296	−37.4%
2014	10.24	1,677,774	2,011,730	−333,956	−19.9%
2015	11.35	1,860,403	1,838,583	21,820	1.2%
2016	11.83	1,940,409	2,233,312	−292,904	−15.1%
2017	12.49	2,050,671	2,545,690	−495,018	−24.1%
2018	15.76	2,592,127	1,971,529	620,599	23.9%
2019	14.26	2,347,030	1,990,885	356,145	15.2%

Table 4.7: Computed club value using Markham's multivariate method compared to the market capitalization of Manchester United for the years 2012-2019. Monetary amounts are given in GBP (thousands).

Additionally, we define the *RT ratio* (see Table A.4 in the Appendix) to provide a relative measure of the difference between the estimated club value and the market capitalization for each year. Presented in Table 4.7, the *RT ratio* ranges from 1.2% to 37.4% in absolute terms. We also define the *Average RT ratio* (see Table A.5 in the Appendix) as a goodness of fit metric for Markham's multivariate method over time, providing a cumulative average of the *RT ratio* for the investigated observations in absolute terms. In this case, the *Average RT ratio* is 18.6%.

Figure 4.1 suggests that Markham's valuation may overestimate the club value compared to the market capitalization of Manchester United. However, this is contradicted by the *RT ratio* suggesting no bias of Markham's valuation overestimating Manchester United's financial valuation, as the ratio takes on positive and negative values an equal amount of times for the investigated observations. This could be further explored by applying the *Average RT ratio* for the positive and the negative *RT ratios*, respectively, where the *Average RT ratio* is 24.1% for the negative values and 13.0% for the positive values. Although this might suggest some bias towards overvaluation, as negative values imply Markham's valuation to be higher than the market capitalization, overvaluation is not consistent over time.

In the Appendix (see Table A.3), the calculations summarized above in Table 4.7 can be found. The historical share prices of Manchester United plc (YF 2022a) and the currency exchange rates for USD/GBP (YF 2022b) at each given date are retrieved from Yahoo Finance. The total number of shares outstanding is retrieved from the annual reports of Manchester United plc (2012-2019), given by the sum of class A and class B shares.

4.3 Application of Results

The results in Section 4.1 show that the best model fit is identified for Model 2 and the Premier League, using an SER approach. For this model, both the intercept and the coefficient are significant on a 0.1% significance level (see Table 4.2) and the null hypothesis (see Section 2.4) can be rejected, implying that sports performance (in terms of league position) can explain club value. Hence, the regression results from this model will be applied to investigate this relationship further. The fitted model looks as follows

$$\widehat{ln(CV)} = 12.016 - 0.027 LeaguePosition, \tag{10}$$

which can be re-written as

$$\widehat{CV} = e^{12.016-0.027 LeaguePosition} = 165,365 \times 0.97^{LeaguePosition}. \tag{11}$$

Using this fitted model, club values in the Premier League can be predicted depending on league position at the end of the season, with the results displayed in Table 4.8. Here, $\widehat{CV}$ gain/loss refers to a club's increase/decrease in estimated club value from gaining/losing one league position with respect to its current position. For instance, moving from third to second place results in an estimated absolute club value gain of GBP (thousands) $4,237$. Likewise, losing one league position from second to third results in an estimated absolute club value loss of GBP (thousands) $4,237$.

For all positions, gaining one league position results in a 2.8% estimated club value gain, and losing one league position results in a 2.7% estimated club value loss. Hence, the relative gain/loss in the estimated club value is constant, but the absolute gain/loss club value is greater for the top clubs due to the exponential relationship. From a monetary perspective, this implies that stakes are higher for the top clubs with respect to gaining/losing club value based on sports performance.

Table A.2 in the Appendix provides a descending list of computed club values for all clubs in the data set. In 2018, the predicted club value of GBP (thousands) $160,889$ from winning the Premier League (see Table 4.8) would only score 17th place on this list. However, this looks reasonable from observing the model's regression plot (see Figure A.3b in the Appendix). Likely, this may be due to a majority of club values for teams in the mid/bottom section of the league table taking on relatively small values. As seen in Section 4.1.2, Model 2 for the Premier League does prove an exponential relationship between club value and league position and is considered the key insight from all regression results.

League Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\hat{CV}$ **gain**	-	$4,355$	$4,237$	$4,122$	$4,011$	$3,902$	$3,796$	$3,694$	$3,594$	$3,496$
$\hat{CV}$ **loss**	$-4,355$	$-4,237$	$-4,122$	$-4,011$	$-3,902$	$-3,796$	$-3,694$	$-3,594$	$-3,496$	$-3,402$
$\hat{CV}$	$160,889$	$156,534$	$152,297$	$148,175$	$144,164$	$140,262$	$136,466$	$132,772$	$129,178$	$125,682$
League Position	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
$\hat{CV}$ **gain**	$3,402$	$3,310$	$3,220$	$3,133$	$3,048$	$2,966$	$2,885$	$2,807$	$2,731$	$2,657$
$\hat{CV}$ **loss**	$-3,310$	$-3,220$	$-3,133$	$-3,048$	$-2,966$	$-2,885$	$-2,807$	$-2,731$	$-2,657$	-
$\hat{CV}$	$122,280$	$118,970$	$115,750$	$112,617$	$109,569$	$106,603$	$103,718$	$100,911$	$98,179$	$95,522$

Table 4.8: Predicted club values dependent on league position at the end of the season for the Premier League. Amounts are given in GBP (thousands).

5 Discussion

This section is divided into two parts. Firstly, the results are discussed in relation to the research question stated in Section 1.2, i.e. *Can the financial valuation of football clubs be explained by their sports performance?* Secondly, considerations with respect to validity, reliability, and generalizability of the results are discussed.

5.1 Sports Performance and Valuation

Overall, the results presented in Section 4.1 suggest that the financial valuation of football clubs can be explained by sports performance. Thus, the null hypothesis stated in Section 2.4 can be rejected, implying a positive correlation (negative correlation with respect to league position) between a football club's sports performance and its financial valuation.

However, the results differ slightly for the different data sets and models used. An SLR approach (see Section 4.1.1) yields significant results for the data sets ALL and EPL, while there is no significance for the EFL data set. This suggests that sports performance has high explanatory power on a club's valuation for the Premier League and both leagues in combination. Nevertheless, the scatterplots for the different data sets (see Figure 3.1) show that the patterns differ for the separate leagues. By combining them, an increased number of observations and an expanded range for the independent variable may facilitate establishing a linear relationship for both leagues. Since the club values are vastly different between leagues, it can be questioned whether or not it is appropriate to investigate them in combination. Arguably, such differences originate from the perks of playing in the top division, where Premier League clubs are able to generate more revenue and attract higher attendance. Even though the linear relationship holds, it is evident that a gap in club values between the two leagues exists, where Premier League clubs are better positioned from a valuation perspective.

Using an SER approach yields significant results for all data sets (see Section 4.1.2), especially for both leagues in combination and the Premier League, implying an exponential relationship between sports performance and club valuation. For the data sets ALL and EPL, this results in a better fitted model than a linear model (see Table 4.5), particularly for the Premier League. This is also coherent with the scatterplot in Figure 3.1b, indicating an exponential increase in club value dependent on an improved league position. Additionally, this is consistent with the literature (see Section 2.1), suggesting that clubs seldom make money except for the top clubs, where higher revenue, correlating with top league positions, drives club valuation. As a result, from a club value perspective, the stakes are higher for better teams in the Premier League fighting for the top positions.

Application of results (see Section 4.3) shows, in practice, how sports performance (in terms of league position) can explain club value in the Premier League through an exponential relationship. Here, absolute numbers confirm that clubs fighting for the top positions have more at stake concerning their club valuation. This implies higher incentives for high-performing clubs to increase their club valuation by improving their league position, in addition to their objective of winning the league. Also, there are other incentives for the best clubs to finish in the top positions. For instance, qualifying for European club competitions is important for a club's ability to generate broadcasting and commercial revenue, acquire better players, and increase brand value.

The incentives for clubs in the mid-section of the table are not as strong, neither from a financial nor a sports perspective, since the European competitions are out of reach and the risk of being relegated to The Championship is low. On the other hand, clubs in the bottom section of the table still have the incentive of not being relegated to The Championship, both from a competitive sports perspective and a financial perspective. This creates an interesting setting in the Premier League, where the incentives for teams on different sports performance levels differ. Although, from the exponential relationship, the relative gain/loss in club value explained by sports performance is constant for all league positions. Hence, an increase/decrease in club value from gaining/losing league positions for low-performing clubs could still substantially impact a club's current value. However, this impact is not of the same magnitude in terms of monetary amounts as for the top clubs.

5.2 Evaluation of Results

As mentioned in Section 2.3, there are implications to valuing football clubs and an approximate method is used in terms of Markham's multivariate method. Hence, this may impact the validity of this study. Also, Manchester United is the only publicly listed English football club where the approximated club values can be evaluated in relation to market values (see Section 4.2). Nevertheless, the estimated club values are relatively coherent in this case, which brings some credibility to the multivariate method and the results of this study. However, worth noting is that Manchester United may be considered an outlier, generating a significantly higher club value than all predictions in the best fitted model (see Section 4.3). Also, there are difficulties in confirming the validity of Markham's multivariate method in general since the remaining clubs do not have a market value for comparison.

Moreover, the reliability of this study is considered high. Likely, similar results could be achieved for repeated tests under similar conditions. However, since the approximated club values used as observations are mainly based on accounting figures, deviations of specific line items can strongly influence the estimated club value, which would not affect ”true” market values in a similar manner. For instance, figures from the income statement could be affected by tax planning that reduces the net profit, which in turn decreases the approximated club value. Moreover, such adjustments made for accounting purposes may only occur for some years, resulting in outlier observations. As seen in Section 4.1.5, outliers exist in all models, which affects the results. However, for our best fitted model, elaborated on in Section 4.3, the effect of outliers is not significant based on goodness of fit statistics after applying winsorizing (see Section 4.1.5).

The generalizability of this study is high, although without certainty. The same results may be seen taking a similar approach to investigate this topic in other settings, for instance, by targeting football leagues in other countries. Nonetheless, as stated in Section 2.1, the market values of football clubs tend to correlate with the geographical region's economic performance. This implies that similar findings may depend on a resembling economic environment, compared to England, in the geographical target region.

6 Conclusion

This paper investigates the relationship between English football clubs' sports performance and their financial valuation. Research on this topic is scarce, while a body of research has examined the relationship between football clubs' sports performance and their financial performance with contradicting results. In general, football clubs tend to prioritize sports performance at the cost of financial performance. From a financial valuation perspective, the football industry differs from other sectors due to characteristics that result in main value drivers such as brand image and sporting success rather than profitability. Due to this and the fact that most football clubs are not publicly listed, conventional valuation methods such as discounted cash flow valuation and trading multiples are not appropriate. An alternative approach can be used by implementing Markham's multivariate method for approximate valuation of football clubs, being applicable to any club.

For the data set consisting of 72 clubs across 22 years in the Premier League and The Championship, evidence suggests that the sports performance of English football clubs does possess explanatory power of their financial valuation. However, the results differ depending on the fitted regression model used and the applied data set. Using an exponential regression model, sports performance in terms of league position demonstrates significance in explaining club value for both leagues in combination and the Premier League and The Championship separately. Applying a linear regression model, league position is significant for the Premier League and both leagues combined, but not for The Championship. It is concluded that an exponential regression model applied to the Premier League yields the best fitted model, followed by an exponential regression model applied to both leagues in combination. Applying a multiple regression approach, using league position and goal difference as different measures of sports performance, there are implications with multicollinearity that undermine the results, and the model is deemed not trustworthy.

Using the best fitted model, the application of results shows that an exponential relationship exists between football clubs' financial valuation and their sports performance in the Premier League. Predictions suggest that the sports performance of high-performing clubs can explain a larger impact on their financial valuation in absolute monetary terms. Vice versa, for low-performing clubs, sports performance explains a smaller impact on financial valuation in absolute terms. However, the explained impact in relative terms is constant for all league positions. This suggests a gap in incentives for teams at different levels, where sports performance of top clubs can explain a higher increase/decrease in club valuation. Hence, concerning financial valuation, high-performing clubs have more at stake in addition to their primary objective of achieving sporting success.

These results contribute both to the field of accounting and sports as well as practitioners in the football industry, providing insights on how sports performance can explain financial valuation. In particular, the research conducted suggests that an exponential relationship is superior to a linear relationship for explaining the relationship between sports performance and financial valuation.

Furthermore, after successfully finding that sports performance can explain club valuation in English football, especially in the Premier League, suggestions for further research can be made regarding the same topic in other settings. Firstly, a similar analysis can be made for English football clubs that accounts for sports performance in a European setting by including achievements in UEFA Champions League, UEFA Europa League, and UEFA Europa Conference League. Secondly, a similar approach could investigate if the same conclusions can be drawn for football clubs in other domestic settings. For instance, interesting settings could be the Italian and Portuguese first divisions, Serie A and Primeira Liga, respectively, given that three clubs in each league are publicly listed on their domestic stock exchange. Thus, it also provides opportunities to test Markham's multivariate method more thoroughly in relation to market capitalization, perhaps by applying the $RT\ ratio$ and the $Average\ RT\ ratio$. Finally, it would be interesting if future work is done using variants of Markham's multivariate method to estimate club value, for instance, through the adjusted version mentioned in Section 2.1 that accounts for EBITDA and amortization costs.

7 References

Aglietta, Michel, Wladimir Andreff, and Bastien Drut (Jan. 2010). *Floating European football clubs in the stock market*. EconomicX Working Papers.

Alaminos, David, Ignacio Esteban, and Manuel Fernández-Gámez (Sept. 2020). "Financial Performance Analysis in European Football Clubs". In: *Entropy* 22, p. 1056. DOI: 10.3390/e22091056.

Andon, Paul and Clinton Free (Oct. 2019). "Accounting and the business of sport: past, present and future". In: *Accounting, Auditing Accountability Journal* ahead-of-print. DOI: 10.1108/AAAJ-08-2019-4126.

Ball, Ray and Philip Brown (1968). "An Empirical Evaluation of Accounting Income Numbers". In: *Journal of Accounting Research* 6.2, pp. 159–178. ISSN: 00218456, 1475679X.

Barros, Carlos Pestana and Stephanie Leach (2006). "Performance evaluation of the English Premier Football League with data envelopment analysis". In: *Applied Economics* 38.12, pp. 1449–1458. DOI: 10.1080/00036840500396574.

Berk, J.B. and P. DeMarzo (2016). *Corporate finance*. Always learning. Pearson. ISBN: 9781292160160.

Botoc, Claudiu, Eugen Mihancea, and Alin Molcut (2019). "Football and Stock Market Performance Correlation: Evidence from Italy". In: *Scientific Annals of Economics and Business* 66.4, pp. 525–539. DOI: 10.47743/saeb-2019-0044.

Breusch, T. S. and A. R. Pagan (1979). "A Simple Test for Heteroscedasticity and Random Coefficient Variation". In: *Econometrica* 47.5, pp. 1287–1294. ISSN: 00129682, 14680262.

Cook, R. Dennis (1977). "Detection of Influential Observation in Linear Regression". In: *Technometrics* 19.1, pp. 15–18. ISSN: 00401706.

Cook, R. Dennis (1979). "Influential Observations in Linear Regression". In: *Journal of the American Statistical Association* 74.365, pp. 169–174. ISSN: 01621459.

Cooper, Christine and Joanne Johnston (May 2012). "Vulgate accountability: Insights from the field of football". In: *Accounting, Auditing Accountability Journal* 25, pp. 602–634. DOI: 10.1108/09513571211225060.

Croux, Christophe, Geert Dhaene, and Dirk Hoorelbeke (Jan. 2004). "Robust Standard Errors for Robust Estimators". In: *Katholieke Universiteit Leuven, Open Access publications from Katholieke Universiteit Leuven*.

Dalso, Jakob and Casper Friis Jensen (2020). *Valuation of Football Clubs: A thesis on valuation of Danish football clubs*.

Deloitte (2022). *Football Money League*. URL: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/sports-business-group/deloitte-uk-dfml22.pdf (visited on 01/27/2022).

Demirakos, Efthimios G., Norman C. Strong, and Martin Walker (Dec. 2004). "What valuation models do analysts use?" English. In: *Accounting Horizons* 18.4, pp. 221–240. ISSN: 0888-7993. DOI: 10.2308/acch.2004.18.4.221.

Dimitropoulos, Panagiotis (2014). "Capital structure and corporate governance of soccer clubs: European evidence". In: *Management Research Review* 37.7, pp. 658–678.

EY (2022). *Premier League economic and social impact*. URL: https://resources.premierleague.com/premierleague/document/2022/01/14/02e0ede6-7ed9-4110-80a5-8e9c04fac776/EY-report-on-Premier-League-economic-and-social-impact-Jan-2022.pdf?utm_source=premier-league-website&utm_campaign=website&utm_medium=link (visited on 03/21/2022).

Fabozzi, Frank J., Sergio Focardi, and Caroline Jonas (2017). "Equity Valuation Science, Art, or Craft?" In: *ERN: Initial Public Offerings (IPOs) (Topic)*.

Feltham, Gerald A. and James A. Ohlson (1995). "Valuation and Clean Surplus Accounting for Operating and Financial Activities." In: *Contemporary Accounting Research* 11.2, pp. 689–731. ISSN: 08239150.

Ferri, Luca et al. (2017). "Financial Versus Sports Performance: The Missing Link". In: *International Journal of Business and Management* 12.3, pp. 1–36.

Galariotis, Emilios, Christophe Germain, and Constantin Zopounidis (2018). "A combined methodology for the concurrent evaluation of the business, financial and sports performance of football clubs: the case of France". In: *Annals of Operations Research* 266.1, pp. 589–612. DOI: 10.1007/s10479-017-2631-z.

Ghosh, Dhiren and Andrew Vogt (2012). *Outliers : An Evaluation of Methodologies*.

Imam, Dr Shahed, Richard Barker, and Colin Clubb (2008). "The Use of Valuation Models by UK Investment Analysts". In: *European Accounting Review* 17.3, pp. 503–535. DOI: 10.1080/09638180802016650.

Karuna, Christo (Mar. 2019). “Capital markets research in accounting: Lessons learnt and future implications”. In: *Pacific-Basin Finance Journal* 55. DOI: 10.1016/j.pacfin.2019.03.001.

Klobučník, Michal, Martin Plešivčák, and Milan Vrábeľ (2019). “Football clubs’ sports performance in the context of their market value and GDP in the European Union regions”. In: *Bulletin of Geography. Socio-economic Series* 45.45, pp. 59–74. DOI: 10.1515/19327.

Koller, T., M. Goedhart, and D. Wessels (2015). *Valuation: Measuring and Managing the Value of Companies, University Edition*. Wiley Finance. Wiley. ISBN: 9781118873731.

Lange, David (2022). *Soccer - Statistics Facts*. URL: https://www.statista.com/topics/1595/soccer/#topicHeader__wrapper (visited on 04/04/2022).

Malagila, John K. et al. (2021). “Corporate governance and performance in sports organisations: The case of UK premier leagues”. In: *International Journal of Finance & Economics* 26.2, pp. 2517–2537. DOI: 10.1002/ijfe.1918.

Manchester United plc (2012-2019). *Annual Reports 2012-2019*. URL: https://ir.manutd.com/financial-information/annual-reports/ (visited on 04/16/2022).

Markham, Tom (Mar. 2013). “What is the Optimal Method to Value a Football Club?” In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2238265.

Montgomery, Douglas C., Elizabeth A. Peck, and Geoffrey G. Vining (2006). *Introduction to Linear Regression Analysis*. 4th edition. Wiley & Sons. ISBN: 0471754951.

Newbold, Paul, William L. Carlson, and Betty M. Thorne (2013). *Statisics for Business and Economics*. 8th edition. Always learning. Pearson. ISBN: 9780132745659.

Nielsen (2018). “World Football Report”. In: URL: https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/world-football-report-2018.pdf (visited on 02/23/2022).

Ohlson, James A. (1995). “Earnings, Book Values, and Dividends in Equity Valuation.” In: *Contemporary Accounting Research* 11.2, pp. 661 –687. ISSN: 08239150.

Pappalardo, Luca and Paolo Cintia (May 2017). “Quantifying the relation between performance and success in soccer”. In: *Advances in Complex Systems* 21. DOI: 10.1142/S021952591750014X.

Peeters, Thomas and Stefan Szymanski (Sept. 2013). *Financial fair play in European football*.

Pereira, Tiago (2018). *The Relation of Financial Performance and the Sports Performance in Football Clubs*.

Pinto, Jerald E, Thomas R Robinson, and John D Stowe (2019). "Equity valuation: A survey of professional practice". eng. In: *Review of financial economics* 37.2, pp. 219–233. ISSN: 1058-3300.

PWC (Oct. 2018). *Accounting for typical transactions in the football industry - Issues and solutions under IFRS*. URL: `https://www.pwc.com/gx/en/audit-services/ifrs/publications/ifrs-9/accounting-for-typical-transactions-in-the-football-industry.pdf` (visited on 04/13/2022).

Reifman, Alan and Kristina Garrett (Jan. 2010). "Winsorize". In: *N. J. Salkind (Ed.), Encyclopedia of Research Design*, pp. 1636–1638.

Shvili, Jason (2020). *The Most Popular Sports In The World*. URL: `https://www.worldatlas.com/articles/what-are-the-most-popular-sports-in-the-world.html` (visited on 02/16/2022).

Simone, Luca Di and Davide Zanardi (2020). "On the relationship between sport and financial performances: an empirical investigation". In: *Managerial Finance* 2020.

Skogsvik, Kenth (2002). "A Tutorial on Residual Income Valuation and Value Added Valuation". In: *SSE/EFI Working Paper Series in Business Administration No 1999:10*.

Sloane, Peter James (1971). "The Economics of Professional Football: The Football Club as a Utility Maximiser". In: *Scottish journal of political economy.* 18.2, pp. 121–146. ISSN: 0036-9292.

Szymanski, S. and S. Kuper (2015). *Money and Football: A Soccernomics Guide*. Nation Books. ISBN: 9781568585260.

Szymanski, Stefan (1998). "Why is Manchester United So Successful". In: *Business Strategy Review* 9, pp. 47–54.

Szymanski, Stefan and Ron Smith (1997). "The English Football Industry: profit, performance and industrial structure". In: *International Review of Applied Economics* 11.1, pp. 135–153. DOI: `10.1080/02692179700000008`.

The Esk (Oct. 2017). *The valuing of football clubs*. URL: `https://theesk.org/2017/10/26/the-valuing-of-football-clubs/` (visited on 03/12/2022).

Tiscini, R. and Dello Strologo (2016). "What drives the value of football clubs: An approach based on private and socio-emotional benefits". In: *Corporate Ownership Control*. DOI: `https://doi.org/10.22495/cocv14i1c4art14`.

Transfermarkt (2022). *Attendances*. URL: https://www.transfermarkt.com/championship/besucherzahlen/wettbewerb/GB2/saison_id/2021 (visited on 03/02/2022).

UEFA, Union of European Football Associations (2022). *UEFA Country Coefficients*. URL: https://www.uefa.com/nationalassociations/uefarankings/country/#/yr/2022 (visited on 05/13/2022).

Wang, George C. S. and Charles K. Akabay (1994). "Heteroscedasticity: How to handle in regression modeling". English. In: *The Journal of Business Forecasting Methods Systems* 13.2, p. 11.

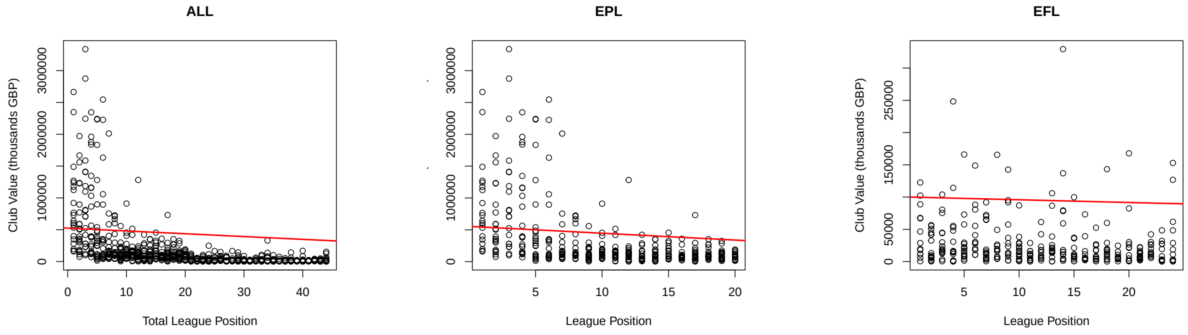
Wood, Robert (2008). *Most Popular Sports Lists Combined.* URL: https://www.topendsports.com/world/lists/popular-sport/analysis.htm (visited on 02/16/2022).

Worldfootball (2022). *Attendance*. URL: https://www.worldfootball.net/attendance/eng-championship-2021-2022/1/ (visited on 03/03/2022).

YF, Yahoo Finance (2022a). *Mancheser Unted plc (MANU): Historical Data*. URL: https://finance.yahoo.com/quote/MANU/history?period1=1344556800&period2=1652572800&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true (visited on 04/20/2022).

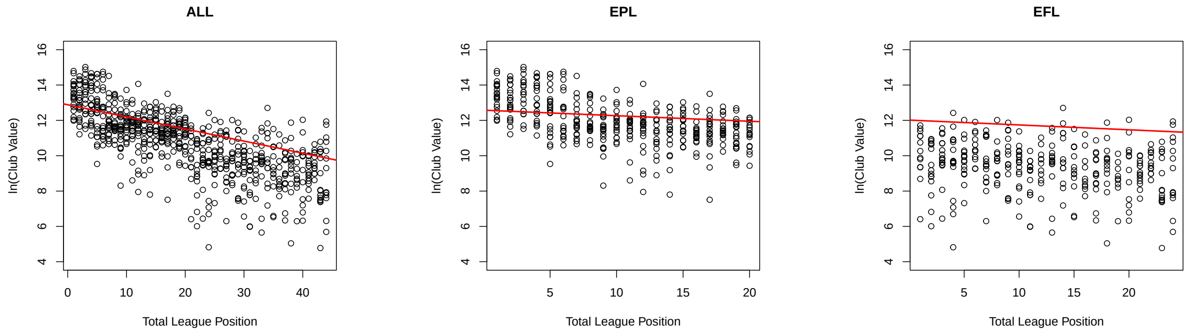
– (2022b). *USD/GBP (GBP=X): Historical Data*. URL: https://finance.yahoo.com/quote/GBP\%3DX/history?period1=1343779200&period2=1652572800&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true (visited on 04/20/2022).

A Appendix



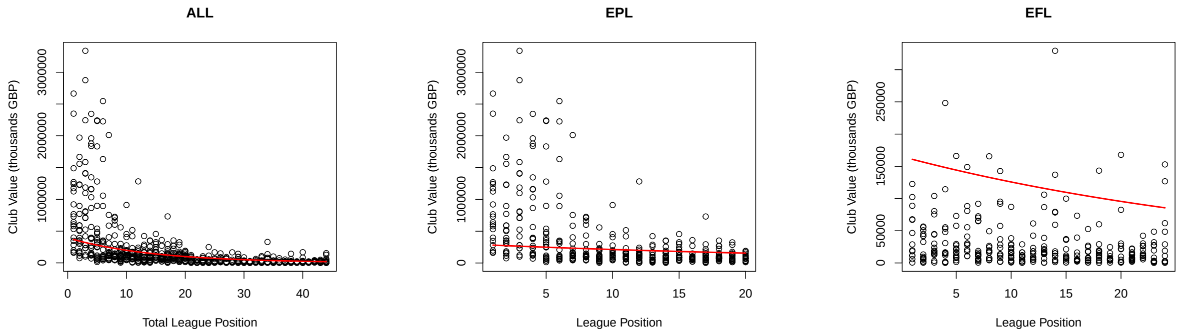
(a) Both leagues (b) Premier League (c) The Championship

Figure A.1: Simple linear regression plots (Model 1).



(a) Both leagues (b) Premier League (c) The Championship

Figure A.2: Simple exponential regression plots (Model 2, Equation 6).



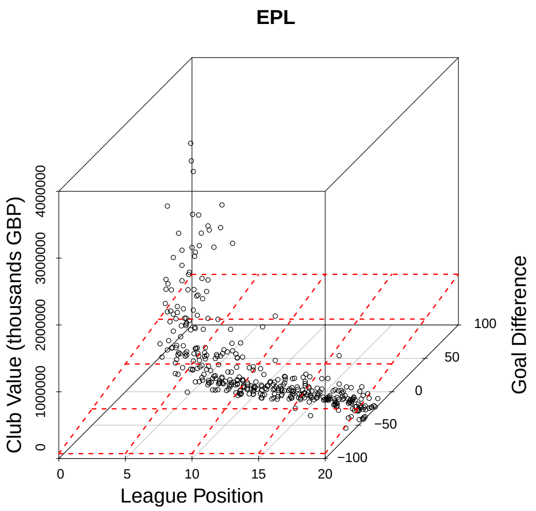
(a) Both leagues (b) Premier League (c) The Championship

Figure A.3: Simple exponential regression plots (Model 2, Equation 7)

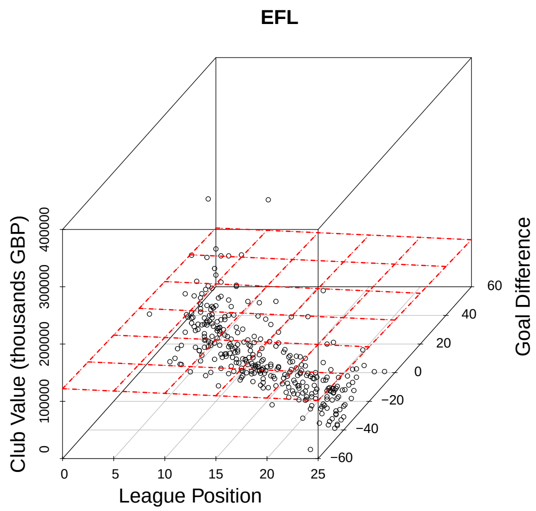
Data set	***EPL***	***EFL***
Dep.var.	CV	CV
Ind. var.	LP, GD	LP, GD
Intercept (β_0)	$414,803^{**}$	$112,315^{***}$
Std.err.	134079.22	16742.41
t-val.	3.0937	6.7084
p-val.	0.0021471	$1.186e-10$
Coefficient (β_1)	144.59	-818.02
Std.err.	5189.55	506.45
t-val.	0.0279	-1.6152
p-val.	0.9777887	0.1074546
Coefficient (β_3)	3405.26^{*}	-162.61
Std.err.	1712.63	203.43
t-val.	1.9883	-0.7994
p-val.	0.0476096	0.4247909
Obs.	386	325
R-sq.	0.6845	0.3581
Adj. R-sq.	0.6286	0.2152
F-stat.	12.23	2.506
p-val.	$< 2.2e-16$	$3.463e-07$

Note: Winsorizing is applied at 5%, Robust standard errors are applied to EPL and EFL
****p<0.001, **p<0.01, *p<0.05*

Table A.1: Multiple linear regression results (Model 3, excl. TP).

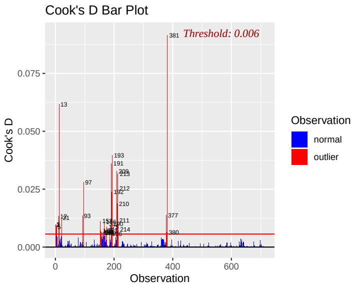


(a) Premier League

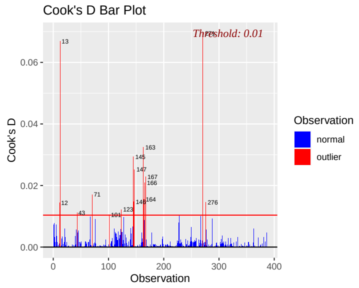


(b) The Championship

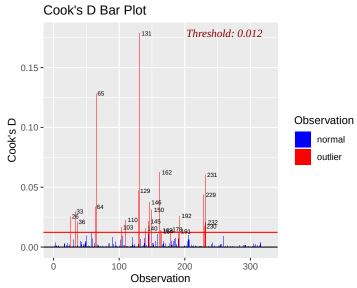
Figure A.4: Multiple linear regression plots (Model 3, excl. TP).



(a) Both leagues

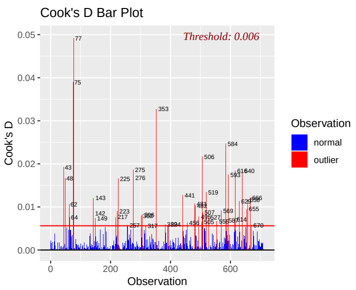


(b) Premier League

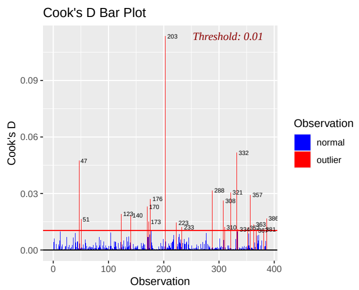


(c) The Championship

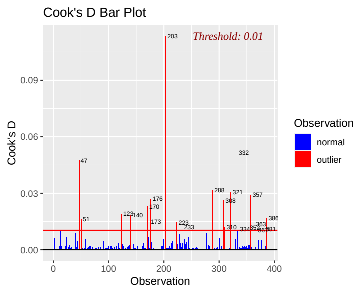
Figure A.5: Cook's Distance plots for Model 1 (simple linear regression)



(a) Both leagues

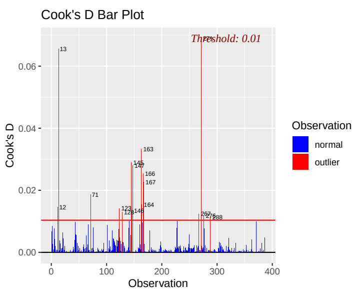


(b) Premier League

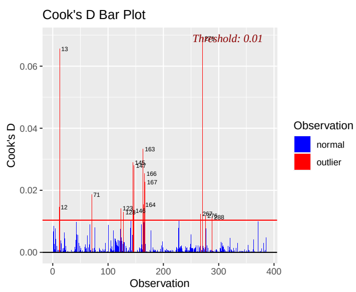


(c) The Championship

Figure A.6: Cook's Distance plots for Model 2 (simple exponential regression)

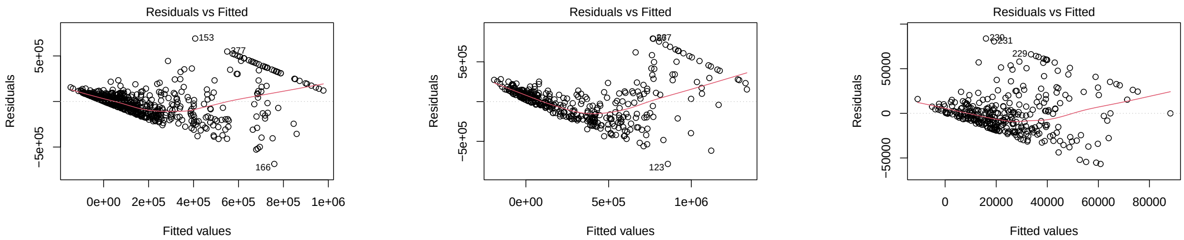


(a) Premier League

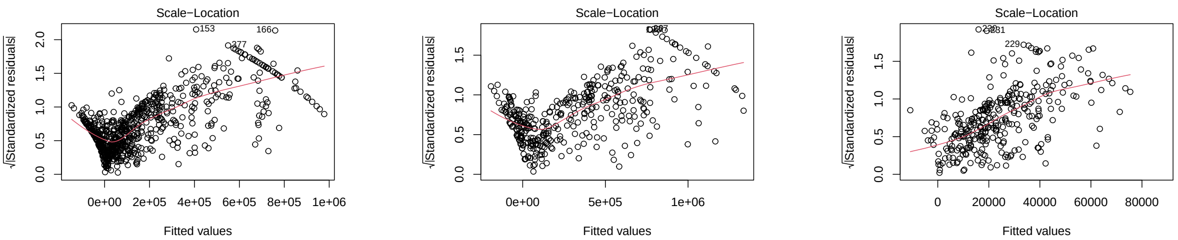


(b) The Championship

Figure A.7: Cook's Distance plots for Model 3 (multiple linear regression)

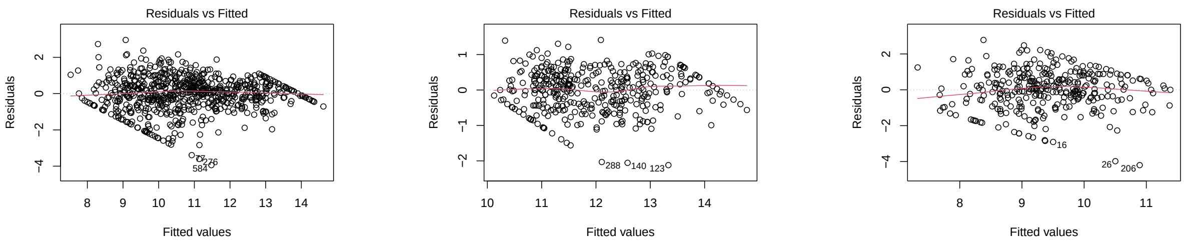


(a) Both leagues (b) Premier League (c) The Championship

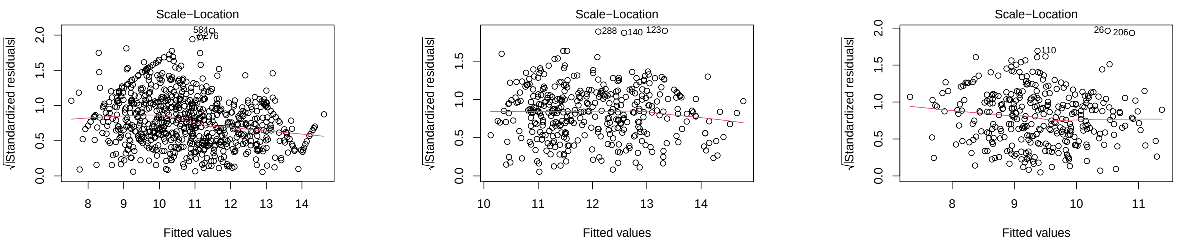


(d) Both leagues (e) Premier League (f) The Championship

Figure A.8: Residual plots for Model 1 - Residuals vs Fitted (upper) and Scale-Location (lower)

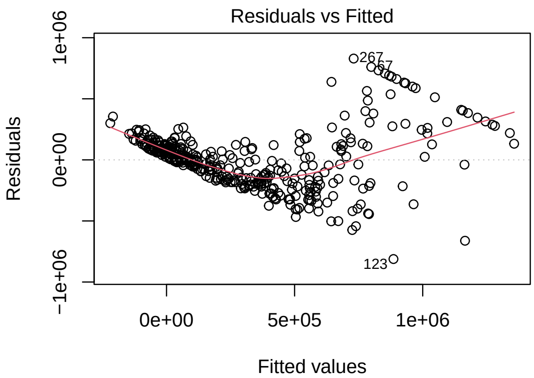


(a) Both leagues (b) Premier League (c) The Championship

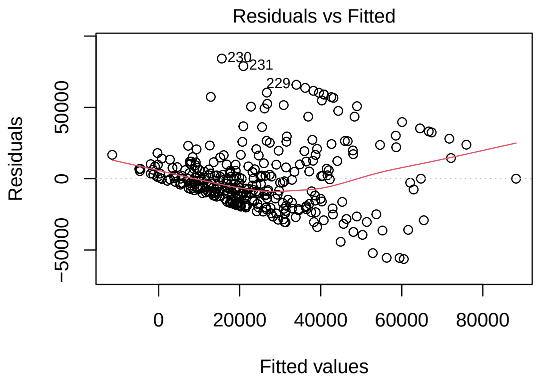


(d) Both leagues (e) Premier League (f) The Championship

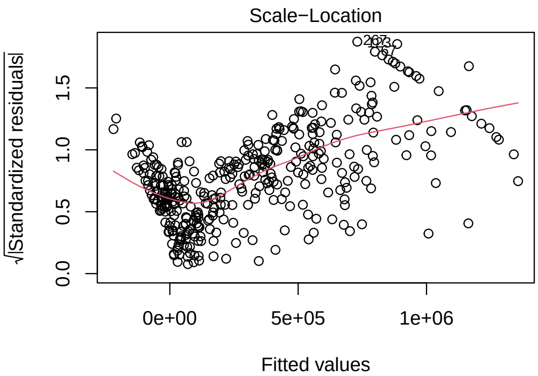
Figure A.9: Residual plots for Model 2 - Residuals vs Fitted (upper) and Scale-Location (lower)



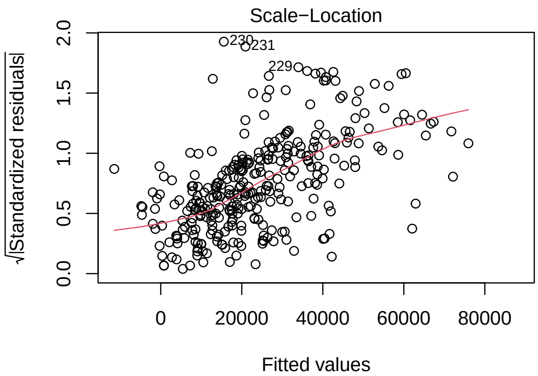
(a) Premier League



(b) The Championship



(c) Premier League



(d) The Championship

Figure A.10: Residual plots for Model 3 - Residuals vs Fitted (upper) and Scale-Location (lower)

	2018				2017				2016			
Club	CV	League	LP	TLP	CV	League	LP	TLP	CV	League	LP	TLP
Tottenham Hotspur	3,337,656	EPL	3	3	1,559,745	EPL	2	2	1,186,656	EPL	3	3
Manchester City	2,664,636	EPL	1	1	2,244,425	EPL	3	3	2,345,329	EPL	4	4
Chelsea FC	2,241,537	EPL	5	5	1,489,075	EPL	1	1	911,079	EPL	10	10
Arsenal FC	2,226,564	EPL	6	6	1,833,007	EPL	5	5	1,234,294	EPL	2	2
Manchester United	1,971,529	EPL	2	2	2,545,690	EPL	6	6	2,233,312	EPL	5	5
Liverpool FC	1,960,640	EPL	4	4	1,032,679	EPL	4	4	600,978	EPL	8	8
Southampton FC	730,862	EPL	17	17	666,236	EPL	8	8	351,201	EPL	6	6
Everton FC	711,828	EPL	8	8	754,616	EPL	7	7	103,552	EPL	11	11
Burnley FC	586,084	EPL	7	7	359,317	EPL	16	16	66,964	EFL	1	21
Leicester City	558,754	EPL	9	9	1,282,213	EPL	12	12	398,281	EPL	1	1
West Ham United	421,860	EPL	13	13	515,419	EPL	11	11	156,233	EPL	7	7
Newcastle United	405,288	EPL	10	10	88,717	EFL	1	21	271,973	EPL	18	18
Swansea City	352,675	EPL	18	18	325,801	EPL	15	15	120,172	EPL	12	12
Norwich City	329,313	EFL	14	34	165,559	EFL	8	28	285,475	EPL	19	19
Huddersfield Town	268,570	EPL	16	16	36	EFL	5	25	N/A	EFL	19	39
Brighton & Hove Albion	208,864	EPL	15	15	3,100	EFL	2	22	N/A	EFL	3	23
Crystal Palace	153,001	EPL	11	11	347,230	EPL	14	14	173,291	EPL	15	15
Derby County	148,887	EFL	6	26	142,624	EFL	9	29	21,019	EFL	5	25
Hull City	143,239	EFL	18	38	265,489	EPL	18	18	N/A	EFL	4	24
AFC Bournemouth	124,889	EPL	12	12	230,457	EPL	9	9	98,229	EPL	16	16
Watford FC	113,134	EPL	14	14	293,372	EPL	17	17	161,972	EPL	13	13
Stoke City	110,260	EPL	19	19	170,067	EPL	13	13	125,995	EPL	9	9
Leeds United	105,974	EFL	13	33	65,214	EFL	7	27	25,691	EFL	13	33
Preston North End	65,222	EFL	7	27	28,890	EFL	11	31	16,219	EFL	11	31
Aston Villa	55,341	EFL	4	24	86,638	EFL	13	33	74,645	EPL	20	20
Sunderland AFC	48,421	EFL	24	44	189,292	EPL	20	20	73,999	EPL	17	17
Sheffield United	29,221	EFL	10	30	-	-	-	-	-	-	-	-
Brentford	28,391	EFL	9	29	38,330	EFL	10	30	23,545	EFL	9	29
Barnsley FC	25,872	EFL	22	42	78,305	EFL	14	34	-	-	-	-
West Bromwich Albion	23,720	EPL	20	20	453,762	EPL	10	10	165,506	EPL	14	14
Middlesbrough FC	22,399	EFL	5	25	151,358	EPL	19	19	26	EFL	2	22
Burton Albion	19,425	EFL	23	43	22,071	EFL	20	40	-	-	-	-
Fulham FC	18,879	EFL	3	23	80,698	EFL	6	26	82,450	EFL	20	40
Nottingham Forest	17,208	EFL	17	37	N/A	EFL	21	41	N/A	EFL	16	36
Birmingham City	10,217	EFL	19	39	N/A	EFL	19	39	8,749	EFL	10	30
Sheffield Wednesday	7,926	EFL	15	35	807	EFL	4	24	10,711	EFL	6	26
Millwall	7,603	EFL	8	28	-	-	-	-	-	-	-	-
Queens Park Rangers	2,237	EFL	16	36	60,002	EFL	18	38	42,380	EFL	12	32
Reading FC	490	EFL	20	40	15,062	EFL	3	23	N/A	EFL	17	37
Cardiff City	407	EFL	2	22	N/A	EFL	12	32	N/A	EFL	8	28
Bristol City	394	EFL	11	31	26,798	EFL	17	37	27	EFL	18	38
Wolverhampton Wanderers	N/A	EFL	1	21	3,884	EFL	15	35	136,944	EFL	14	34
Ipswich Town	N/A	EFL	12	32	N/A	EFL	16	36	N/A	EFL	7	27
Bolton Wanderers	N/A	EFL	21	41	-	-	-	-	152,917	EFL	24	44
Blackburn Rovers	-	-	-	-	N/A	EFL	22	42	N/A	EFL	15	35
Rotherham United	-	-	-	-	10,915	EFL	24	44	12,357	EFL	21	41
Wigan Athletic	-	-	-	-	48,483	EFL	23	43	-	-	-	-
Charlton Athletic	-	-	-	-	-	-	-	-	695	EFL	22	42
Milton Keynes Dons	-	-	-	-	-	-	-	-	1,814	EFL	23	43

CV=Club Value, LP=League Position, TLP=Total League Position

Table A.2: Computed club values in GBP (thousands) for all clubs in Premier League and The Championship for the years 2016-2018, in descending order according to their club value in 2018.

Date	2012-08-10	2013-06-30	2014-06-30	2015-06-30	2016-06-30	2017-06-30	2018-06-30	2019-06-30
Outstanding Shares	155,352,366	163,825,595	163,777,957	163,873,074	164,025,280	164,194,754	164,526,390	164,570,967
Share Price (USD)	14.00	15.92	17.45	17.86	15.93	16.25	20.60	18.08
Market Cap (USD)	2,174,933,124	2,608,103,472	2,857,925,513	2,926,773,266	2,612,922,710	2,668,164,753	3,389,243,634	2,975,443,083
USD/GBP	0.63975	0.65538	0.58706	0.63565	0.74262	0.76857	0.76481	0.78880
Share Price (GBP)	8.96	10.43	10.24	11.35	11.83	12.49	15.76	14.26
Market Cap (GBP)	1,391,413,466	1,709,298,854	1,677,773,752	1,860,403,426	1,940,408,663	2,050,671,384	2,592,127,424	2,347,029,504
Markham Valuation (GBP)	1,227,979,285	2,348,594,858	2,011,730,090	1,838,583,355	2,233,312,223	2,545,689,743	1,971,528,640	1,990,884,830
Difference (GBP)	163,434,181	-639,296,004	-333,956,338	21,820,071	-292,903,560	-495,018,359	620,598,784	356,144,674
RT ratio	11.7%	-37.4%	-19.9%	1.2%	-15.1%	-24.1%	23.94%	15.2%

Table A.3: Complete calculation of values for evaluation of Markham's multivariate model.

Definition: Rylander-Ternqvist ratio

$$RT\ ratio\ =\ \frac{Market\ Capitalization_{i,t}\ -\ Markham\ Valuation_{i,t}}{Market\ Capitalization_{i,t}} \tag{12}$$

for club i at time t.

Table A.4: Rylander-Ternqvist ratio (*RT ratio*), created for this thesis as a simple goodness of fit measurement for Markham's multivariate method's approximated valuation in relation to market capitalization.

Definition: Average Rylander-Ternqvist ratio

$$RT_{avg}\ ratio\ =\ \frac{\sum_{t=a}^{b} \frac{|\ Market\ Capitalization_{i,t}\ -\ Markham\ Valuation_{i,t}\ |}{Market\ Capitalization_{i,t}}}{b-a} \tag{13}$$

for club i during the time period $t\ \epsilon\ [a, b]$

Table A.5: Average Rylander-Ternqvist ratio (*Average RT ratio*), created for this thesis as a cumulative average measure of the RT ratio to summarize the goodness of fit for Markham's multivariate method's approximated valuation in relation to market capitalization.